Review of Industrial Organization (2020) 56:489–513
https://doi.org/10.1007/s11151-019-09706-5

Movie Industry Demand and Theater Availability

Tin Cheuk Leung[1] · Shi Qi[2] · Jia Yuan[3]

Published online: 10 June 2019

Abstract
Consumers can only choose to see a movie if it is available in theaters. Explicitly taking into account movie theater availability, we estimate a structural model of movie demand with the use of U.S. movie data from 1995 to 2017. Estimation results indicate that the impact of theater availability on movie demand is both statistically and economically significant. We also find that movie budget predictions based on the model that incorporates theater availability is more consistent with the data, while the model that ignores theater availability on average over-predict production budgets.

Keywords Movie demand · Theater availability · Demand estimation

1 Introduction

Consumers can choose to see a movie only if it is available in theaters; hence theater availability is an important determinant of movie box-office performances. For example, two movies—*Everyone Says I Love You* and *My Best Friend's Wedding*—share the same release year (1997), star actress (Julia Roberts), and genre (Romantic Comedy), as well as a similar storyline. Despite having a less positive review,[1] the latter was released in far more theaters—2134—compared to 268 for the former

[1] *Everyone Says I Love You* has a rating of 6.8/10 on IMDb, compared to 6.3/10 for *My Best Friend's Wedding*.

✉ Tin Cheuk Leung
leungtc@wfu.edu

Shi Qi
sqi01@wm.edu

Jia Yuan
jiayuan@umac.mo

1 Wake Forest University, Box 7505, Winston-Salem, NC 27109, USA

2 College of William and Mary, Williamsburg, VA, USA

3 University of Macau, Zhuhai, China

in the opening weeks and had a significantly higher box-office revenue—$127 million—compared to $10 million. Through its impact on movie demand, theater availability could significantly affect movie production and investment decisions.[2]

To study the impact of theater availability on movie demand, we construct a structural model that explicitly incorporates theater availability. We estimate the demand model with the use of U.S. movie data from 1995 to 2017. Estimation results indicate that the impact of theater availability on movie demand is both statistically and economically significant. We find that ignoring theater availability can significantly bias the estimated impact of movie production budgets on box-office performances. The demand model that ignores theater availability would mis-attribute a movie's box-office success largely to its perceived quality. This bias can inflate the importance of movie production budgets, which are used by movie studios to increase movie qualities.

To quantify the importance of theater availability, we simulate movie studios' budget decisions based on the estimated demand models. The predicted budgets—with and without considering theater availability—are then compared to the actual movies budgets in the data. We find that the model that takes into account theater availability is more consistent with the data, which indicates that movie studios are mindful of the impacts of theater availability in their budget decisions. Furthermore, the model that ignores theater availability would over-predict the production budget of an average movie by $111 million.

Our paper adds to the growing list of papers studying movie industry demands, including De Vany and Walls (1996, 1997), Einav (2007), Moul (2007a, b), Moretti (2011), Ferreira et al. (2012), and Dalton and Leung (2017). Our structural demand model is akin to that of Einav (2007). The identification assumptions and estimation techniques are similar to those in Moul (2007b). The production budget prediction model is similar to that of Ferreira et al. (2012).

2 Theater Availability

A consumer's choice of movie is bounded by what is available in theaters. A movie's theater availability depends upon theater owners' exhibition decisions. Exhibition decisions, according to Filson et al. (2005) who study movie theater contracts, are completely up to individual theater owners, and can be adjusted according to past movie performances.

To illustrate the mechanism through which theater availability can affect movie demand, consider two hypothetical towns: A and B; each has only one theater. The theater in town A can play only one movie, while the theater in town B can play two movies at the same time. The two towns are otherwise identical. The townspeople can only go to their respective in-town theater to watch movies. Suppose two movies—one of high quality (H) and one of low quality (L)—are released in the same

[2] Theatrical box-office revenue is the largest income source for movie studio, see https://stephenfollows.com/how-movies-make-money-hollywood-blockbusters/.

week. Town A's theater owner plays only movie H, because the high quality movie draws a larger audience. In Town B, both movies are played. Movie L can draw audiences when played in Town B, but it does not enter consumers' considerations in Town A. Therefore, the difference in the two movies' box-office revenues is not only a result of their quality difference, but also reflects the different constraints of theater availability.[3]

Movies' shares of theaters are used to proxy theater availability in the ensuing analysis. A movie's share of theaters in a particular week is defined as the ratio of the number of theaters that play the movie in that week to the total number of theaters in the market.

A few caveats are worth noting here: First, we assume that every theater has the same number of screens. Due to data limits, we cannot account for variations of theaters in size and number of screens. Second, a movie is assumed to have the same number of showings per theater for any week in which it is available. It is certainly possible that a movie has more showings per theater in the beginning than at the end of its release. Third, we cannot differentiate the size and quality of screens. Sometimes, a movie is played on screens that can accommodate larger audiences in the beginning of its release, but relegated to smaller screens in later weeks.

As a consequence of the second and third assumptions, the decline in a movie's box-office revenue over time due to a loss of theater availability can be misattributed to consumers losing interest in the movie. Therefore, our analysis tends to underestimate the importance of theater availability because of these assumptions.

In addition, we are implicitly assuming that the changes in theater shares are proportional across different geographic locations. This assumption is reasonable as long as all movie distributors and exhibitors across the country follow profit maximization.

3 Data

The main data source is *The Numbers*, which is a movie industry data website. The weekly data cover all movies that were released in the United States between Fri. Dec. 30, 1994, and Fri. May 19, 2017. For each movie released, the data report its official release date, weekly box-office revenues, weekly number of theaters, as well as information on movie distributor, genre, and estimated production budget. *The Numbers* also reports the total annual admissions in the industry.

[3] A related concept to theater availability is the concept of "stock-out". For example, Mortimer (2008) finds that video rental stores often experience inventory problems, and some movie titles are unavailable to consumers because they are out of stock. An equivalent phenomenon in movie theaters would be sold-out performances of blockbuster movies. Under the assumption that every theater screen's seating capacity is at least 100 people, and each screen has at least 3 showings a day, we find no movie in our data reaches more than 55% of the total aggregate seating capacity in any week. It is still possible that the sold-out performances happen in opening nights or are restricted to certain local theaters. However, because our movie data are weekly national level data, we do not have a reliable way to identify sold-out performances.

Other industry aggregate data are obtained from several sources. Average movie ticket prices, total numbers of U.S. movie theaters and screens by year are obtained from the *Encyclopedia of Exhibition* of the National Association of Theater Owners. Annual population figures are obtained from the U.S. Census. Linear interpolations are used to estimate ticket prices, aggregate numbers of theaters and screens, as well as the U.S. population for each week in the sample period.

The initial data sample comprises 10,787 movie titles. Some observations are dropped in the analysis. First, we drop small movies that are never in wide-release. Wide-release refers to a movie that played in more than 600 theaters for at least one week during its theatrical run. We follow the literature in making this restriction, because limited-release movies have very different box-office performances and release patterns.

Second, movies are considered only for their first ten full weeks of release, and all observations thereafter are dropped.[4] For a vast majority of movies, box-office revenues become very small after the tenth week in release: They average only 4.3% of their opening-week earnings.

In addition, we drop movies that do not have at least three consecutive weeks of observations. The final data contain 3283 movie titles and 30,432 movie-week observations. These movie observations account for 92.4% of the total box-office revenues in the raw data. Table 1 presents general industry trends over the sample period.

Overall, an average of 147 movies were widely released in the industry each year. We use in the analysis the national average ticket prices. Compared to the Consumer Price Index consisting "of other recreational services", the average movie ticket price increased starting in the late 2000s. Meanwhile, the annual admissions per capita decreased during the same period compared to the years before 2010. The annual admission per capita is the ratio of total annual admissions (total box-office revenue divided by average ticket price) and the U.S. population.

3.1 Summary Statistics

Table 2 summarizes approximated admissions per released movie per week in the sample from 1995 to 2016.[5] Per movie admissions are approximated using the ratio of box-office revenues and the national average ticket prices.[6]

The table is organized by a movie's release year. For example, if released on December 31, 2010, a movie would be counted in 2010, even if most of its theatrical run is in 2011. Columns two to five show the mean, median, minimum, and

[4] In our data, about 13% of the movies were released on a day other than Fridays. For those movies, we used the eleventh week as the last full week.

[5] The sample covers only a fraction of weeks in the years 1994 and 2017; consequently they are excluded from the tables of summary statistics.

[6] The movie prices used are the national annual average prices. We follow Einav (2007) and use linear interpolation to obtain weekly average prices. The data do not have price variations by geographic areas or age groups. Therefore, in the calculation of weekly admissions per movie, we are implicitly assuming that all movies have the same proportional exposures to different geographic regions and age groups.

Table 1 Industry trends

Year	# of movies released	Avg. ticket price	Total box-office revenues ($ billions)	Annual admissions (per capita)
1995	139	$7.37	8.3	4.25
1996	131	$7.26	8.2	4.22
1997	135	$7.34	8.9	4.48
1998	134	$7.24	8.7	4.37
1999	143	$7.42	9.6	4.65
2000	145	$7.61	9.8	4.57
2001	137	$7.68	10.0	4.58
2002	148	$7.64	10.5	4.79
2003	141	$7.68	10.6	4.77
2004	151	$7.76	10.8	4.77
2005	146	$7.78	9.8	4.26
2006	171	$7.65	10.2	4.47
2007	178	$7.86	10.5	4.44
2008	162	$8.02	10.2	4.19
2009	149	$8.36	11.5	4.50
2010	133	$8.72	10.7	4.00
2011	142	$8.70	10.4	3.83
2012	140	$8.52	10.8	4.06
2013	147	$8.61	10.9	4.01
2014	144	$8.58	10.4	3.81
2015	149	$8.61	10.9	3.94
2016	160	$8.65	10.9	3.90

Ticket prices and revenues are in December 2016 U.S. dollars, adjusted using the CPI for "Other Recreation Services"

maximum of box-office revenues by release year. Weekly admissions vary greatly among movies: They range from less than a million to more than 100 million. A consistent feature is that the median is much smaller than the mean in any given year. This indicates that the box-office admissions across movies are highly skewed, such that a few blockbusters earn a large portion of the total admission in any given week.

Average weekly admissions by week are presented in Fig. 1, in which all of the major holidays are labeled. The left panel of Fig. 1 shows the industry seasonal pattern. The peak seasons are in the summer and during the winter holidays. The summer peak season extends from Memorial Day to Labor Day. The winter holiday season starts at Thanksgiving and ends roughly at New Year. Spring and autumn are traditionally the low seasons.

Examining weekly admissions in three different time periods, 1995–2000, 2001–2009, and 2010–2016, the right panel of Fig. 1 shows that the underlying seasonal patterns remain largely the same over the sample period.

Table 2 Weekly admission per movie (in millions of tickets sold)

Year	Mean	Median	Max	Min
1995	8.10	5.35	41.38	0.61
1996	8.63	4.98	55.01	0.93
1997	8.99	6.20	94.02	0.73
1998	8.98	5.40	39.44	0.25
1999	9.04	5.32	69.96	0.38
2000	8.86	6.04	48.15	0.20
2001	9.49	5.67	55.21	0.55
2002	9.30	5.69	69.37	0.52
2003	9.77	6.80	60.02	0.27
2004	9.23	5.73	69.43	0.43
2005	8.60	5.48	58.74	0.30
2006	7.77	5.36	63.94	0.47
2007	7.48	4.06	48.74	0.14
2008	7.84	4.77	73.12	0.31
2009	9.25	5.08	95.00	0.22
2010	9.26	6.11	51.48	0.36
2011	8.38	5.02	47.79	0.18
2012	9.08	5.77	77.19	0.13
2013	8.60	4.93	51.88	0.33
2014	8.42	5.37	41.04	0.08
2015	8.46	3.90	109.96	0.12
2016	7.84	4.02	61.23	0.07

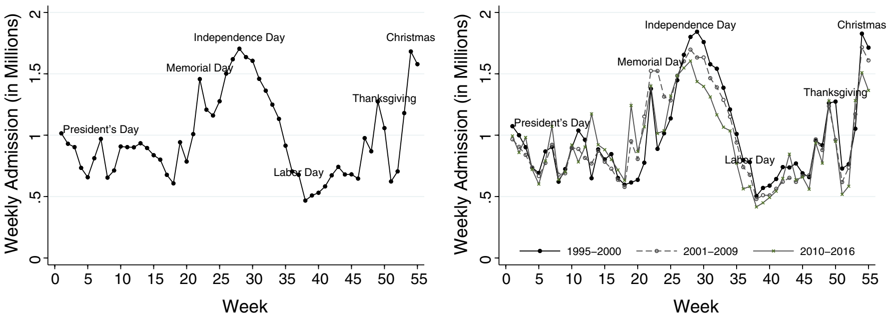

Fig. 1 Average weekly admission (in millions)

The data are weekly in nature. Following the previous literature, weeks are adjusted so that major holidays (which are labeled in Fig. 1) fall in the same week across all years. For example, Labor Day and Thanksgiving are always assigned the same week numbering: Week 37 and 50, respectively. However, depending on the year, there might be 11 or 12 weekends between the two holidays. Week 43 is included as a "filler" in those years with 12 weeks between the two holidays. In

Table 3 Weekly numbers of theaters and screens per movie

Year	Weekly number of theaters per movie				Total of theaters #	Total # of screens	Avg. screens per theater
	Mean	Median	Max	Min			
1995	1074.4	969	2893	8	7744	27,843	3.60
1996	1208.9	1131	3012	29	7798	29,731	3.81
1997	1190.4	1019	3565	19	7480	31,865	4.26
1998	1254.8	1085	3310	21	7418	34,168	4.61
1999	1207.0	944	3411	2	7477	35,506	4.75
2000	1311.3	1175	3669	11	6992	35,650	5.10
2001	1320.3	1103	3715	12	6253	35,688	5.71
2002	1244.2	901	3876	2	6144	36,379	5.92
2003	1307.0	968	3749	1	6100	36,435	5.97
2004	1272.1	867	4223	3	6031	37,131	6.16
2005	1342.9	1002	4142	1	6114	37,688	6.16
2006	1238.3	781	4133	1	5939	38,415	6.47
2007	1227.7	795	4362	2	5928	38,794	6.54
2008	1278.5	763	4366	2	5786	39,233	6.78
2009	1382.3	985	4393	1	5942	39,520	6.65
2010	1429.9	971	4468	2	5773	39,580	6.86
2011	1358.8	819	4375	5	5697	39,662	6.96
2012	1355.9	840	4404	6	5683	39,956	7.03
2013	1379.8	905	4253	7	5719	40,006	7.00
2014	1374.2	830	4324	1	5856	40,024	6.83
2015	1322.6	791	4311	2	5833	40,174	6.89
2016	1303.2	658	4381	1	5821	38,834	6.67

essence, Week 43 would be missing in those years with 11 weeks between the two holidays. The filler weeks are Weeks 13, 23, 33, 43, 53 and 55, which do not appear in all years.[7]

Table 3 summarizes theater and screen statistics per released movie per week. Columns two to five show the mean, median, minimum, and maximum of theaters by release year. The means and median numbers of theaters are largely flat throughout the sample years. The maximum number is increasing from 1995 to the mid-2000s. After 2007, the trend is largely flat. In the data, most wide-release movies reach more than 2000 theaters. However, as the minimum numbers show, these wide-release movies can be in very few theaters at the end of their theatrical runs.

Table 3 also shows the industry aggregate numbers of theaters and screens. While the total number of screens increases substantially, the total number of theaters decreases over the sample period. This is likely due to a major shift from many small single screen theaters to fewer larger multiplex theaters during the same time

[7] Every year has 52 weeks with some filler weeks missing. For example, the year 2015 has only filler weeks 23, 33 and 53, while the year 2014 has only filler weeks 23, 43 and 55.

Table 4 Movies production budgets

Year	Production budget per movie (in millions of Dec. 2016 dollars)			
	Mean	Median	Max	Min
1995	86.27	71.46	367.17	7.46
1996	81.73	80.99	202.77	5.05
1997	99.50	86.50	387.62	5.90
1998	81.78	57.59	429.89	0.19
1999	68.62	56.18	318.91	1.09
2000	69.15	57.92	218.56	1.69
2001	65.90	49.18	245.15	2.67
2002	65.04	53.94	218.27	1.88
2003	70.81	56.78	255.41	0.44
2004	67.85	48.23	290.60	0.58
2005	67.95	52.20	287.41	1.40
2006	59.14	41.52	318.35	2.72
2007	63.10	35.13	399.49	1.00
2008	59.67	38.25	286.30	0.63
2009	65.31	44.55	504.17	0.54
2010	68.47	43.40	294.54	2.06
2011	63.04	44.83	281.14	1.69
2012	62.81	39.11	308.08	1.12
2013	62.55	35.59	303.83	0.99
2014	55.62	30.60	272.10	0.22
2015	55.11	31.80	355.75	0.11
2016	53.87	30.23	262.03	0.94

Production budgets are in December 2016 U.S. dollars, adjusted using the CPI for "Other Recreation Services"

period. As a result, the average number of screens per theater almost doubled from 3.60 in 1995 to 6.67 in 2016.

Table 4 shows the mean, median, minimum, and maximum of inflation-adjusted movie production budgets by release year. The mean budget decreases in the late 1990s, then there is no discernible trend for the rest of the sample years. The difference between the mean and median has widened over the sample periods, which could indicate that movie producers have shifted budgets to a few blockbuster movies.

Furthermore, an average movie's length of theatrical run changes little over the sample years. For example, 64.0% of all movies released in 1995 have theatrical runs of 10 or more weeks, while the corresponding number in 2016 is 64.4%. The average of the entire data sample is 67.8%. This shows that movies do not stay in theaters longer despite the increase in the total number of screens.

Measuring the impact of theater availability on movie box-office performance is an empirical exercise that is undertaken in Sect. 4. Used in the empirical model,

Table 5 Movie share variables

Variable	Description	Mean	Median	Max	Min
s_{jt}/s_{ot}	Relative market share ratio	0.0035	0.0009	0.1775	0.0000
$s_{jt}/(1-s_{ot})$	Within-industry market share	0.0384	0.0105	0.8197	0.0000
m_{jt}	Theater share (%)	22.48	15.62	83.31	0.02

several important variables are directly constructed from the data. First, movie *j*'s market share in week $t - s_{jt}$—is the ratio of the movie's weekly admission (box-office revenue divided by ticket price) to the U.S. population, which is assumed to be the overall potential market size.[8] The fraction of population not going to theaters in week t is $s_{ot} = 1 - \sum_j s_{jt}$.[9] Second, movie *j*'s share of theaters in week $t - m_{jt} = a_{jt}/A_{jt}$—is the ratio of movie *j*'s number of theaters (a_{jt}) to the total number of theaters, A_t. In the ensuing analysis in Sect. 4, the logarithm of s_{jt}/s_{ot} is the left-hand-side variable, and the logarithms of both the within-industry market share, $s_{jt}/(1-s_{ot})$, and the theater share, m_{jt}, are the right-hand-side variables. Table 5 summarizes these variables.

4 Model

This section provides the details of a structural model. The model includes both the demand and supply sides of the movie industry. The demand model is similar to that in Einav (2007) and Moul (2007b), and includes both the movie quality and market-expansion effects. On the supply side, movie producers' investment decisions are modeled taking into account the impact of production budget on demand. On both the demand and supply sides, the model incorporates theater availability. In addition, we discuss the identification assumptions and instrumental variables.

4.1 Demand Specification

A nested logit demand is used to model an individual consumers' movie-going decisions. The utility of consumer *i* from watching movie *j* in week *t* is

$$u_{ijt} = \theta_j - \lambda(t - r_j) + \delta \ln(m_{jt}) + \xi_{jt} + \varsigma_{it} + (1-\sigma)\varepsilon_{ijt}, \qquad (1)$$

[8] Alternatively, we can use weekly revenues to calculate market shares. In this case, the potential market size is calculated by multiplying the total U.S. population and the average ticket price. This method of calculation would yield exactly the same market shares. In addition, the calculation of admissions does not depend on the method of inflation adjustment, because the numerator and denominator (weekly box-office revenues and average ticket prices) would both be adjusted by the same inflation measure.

[9] Implicitly, we are assuming that a potential audience goes to at most a movie per week. This assumption is reasonable because the annual admission per capita is only about 4, as shown in Table 1.

in which: θ_j is a movie's perceived quality; r_j is movie j's initial week of release; $t - r_j$ is movie j's numbers of week-in-release in week t; m_{jt} is movie j's share of theaters; ξ_{jt} is the unobserved propensity to like movie j in week t; and $\varsigma_{it} + (1-\sigma)\varepsilon_{ijt}$ is an idiosyncratic taste shock.

The perceived quality of movie $j - \theta_j$—is invariant across different consumers and weeks. In turn, θ_j depends on movie j's: production budget B_j; genre g; release year y; and a random effect χ_j.

$$\theta_j = \alpha + \beta \ln(B_j) + \mu_g + \psi_y + \chi_j. \tag{2}$$

In a given week, consumer i can choose an outside good (good 0), which determines the consumer's propensity to stay away from movie theaters. Consumer utility of staying away is

$$u_{i0t} = -\tau_t + \varsigma'_{it} + (1-\sigma)\varepsilon_{i0t}, \tag{3}$$

where τ_t is the weekly fixed effect.

The nested nature of the logit demand depends on the idiosyncratic taste shock, $\varsigma_{it} + (1-\sigma)\varepsilon_{ijt}$. The taste shock component, ς_{it}, is the same across all movies, but can be different from the unobserved propensity to choose the outside option, ς'_{it}. We assume that ε_{ijt} is an independent and identically distributed extreme value random variable. The sum $\varsigma_{it} + (1-\sigma)\varepsilon_{ijt}$ is also extreme value distributed. Parameter $\sigma \in [0,1]$ captures the relative importance between these two taste shock components and measures the substitutability between movies and the outside option.

Following Berry (1994), the nested logit predicted market share of movie j in week t is:

$$s_{jt} = \frac{\exp\left(\frac{\theta_j - \lambda(t-r_j) + \tau_t + \delta \cdot \ln(m_{jt}) + \xi_{jt}}{1-\sigma}\right)}{D_t^{\sigma} + D_t}, \tag{4}$$

where

$$D_t = \sum_{k \in J_t} \exp\left(\frac{\theta_k - \lambda(t-r_k) + \tau_t + \delta \ln(m_{kt}) + \xi_{kt}}{1-\sigma}\right).$$

Here, J_t is the set of all available movies in theaters in week t. If we rearrange Eq. (4), we obtain

$$\begin{aligned} \ln(s_{jt}) - \ln(s_{0t}) &= \theta_j - \lambda(t-r_j) + \tau_t + \delta \ln(m_{jt}) + \sigma \ln\left(\frac{s_{jt}}{1-s_{0t}}\right) + \xi_{jt}, \\ &= \alpha + \beta \ln(B_j) + \mu_g + \psi_y + \chi_j - \lambda(t-r_j) + \tau_t + \delta \ln(m_{jt}) \\ &\quad + \sigma \ln\left(\frac{s_{jt}}{1-s_{0t}}\right) + \xi_{jt}. \end{aligned} \tag{5}$$

We show the derivation of both Eqs. (4) and (5) in "Appendix 1". The term s_{0t} enters into both sides of the regression Eq. (5), and the with-in market share—$\frac{s_{jt}}{1-s_{0t}}$—is endogenous. We use an instrumental variable approach to address this problem, which is discussed in Sect. 4.2.

All demand parameters are estimated using Eq. (5). Here, all movies have a baseline quality α, upon which the quality increases with the production budget B_j. The marginal return to budget investment, β, is expected to be positive. The logarithm is used to capture the decreasing return to budget investment. Both genre (μ_g) and release-year (ψ_y) of a movie can also affect a movie's perceived quality.[10] In addition, the random effect term, χ_j, is added to control for unobserved heterogeneity in movie quality, which is due to possible differences in screenplays or concept-ideas.

The rate of decay in a movie-goer's propensity to watch a movie after its release is captured by the parameter λ. We assume that the rate of decay is common across all movies and time periods. The weekly fixed effect—τ_t—captures the underlying seasonality in the movie industry.

Our model incorporates theater availability as an important determinant of a movie's market share as in Moul (2007b). A movie can be considered by consumers only if it is available in theaters.

We use the concept of "consideration set" that is commonly used in the marketing literature.[11] The consideration set is made up of the choices that are seriously considered by a consumer in her purchase decisions. If a theater ceases to show a movie, then the movie is eliminated from the consideration sets of the theater's customers. Therefore, the more theaters to which a movie is released, the higher is the probability that the movie is in consumers' consideration sets. A movie's consideration probability is the probability that it is included in consumers' consideration sets. We assume that the consideration probability of movie j in week t depends on its share of theaters m_{jt}. We use the logarithm form, $\ln(m_{jt})$, because of the decreasing marginal return to adding theaters due to overlaps in market exposures. Parameter δ is the intensity parameter on theater availability.[12]

The parameter σ determines the "market-expansion" effect. A high quality movie can attract consumers who would otherwise stay away from theaters. In doing so, the movie expands the movie-going consumer base. If $\sigma = 1$, then the outside good and all of the movies have no substitutability. This means that a movie can only expand its market share at the expense of other movies. If $\sigma = 0$, the model is a simple logit

[10] The annual dummy variable (ψ_y) captures aggregate changes in the market, including changes in average ticket prices, national incomes, macroeconomic business cycles, etc. We do not use prices directly in the estimation, because only average national ticket prices are available in the data. This means that the aggregate price effect cannot be separately identified from the effect that is due to other possible aggregate factors. In addition, because individual movies in general do not compete on price margins: Conditional on everything else being the same, an audience pays the same price to see any available movies in the same theater. We do not expect unobserved idiosyncratic demand shocks would cause endogeneity bias on the estimation of the annual fixed effects.

[11] See more details in Fotheringham (1988), Bronnenberg and Vanhonacker (1996), and Wu and Rangaswamy (2003).

[12] The consumer choice probability in Eq. (4) has the theater availability directly entering consumers' utility function. This is equivalent to defining separately a consideration probability $\pi_{jt} = (m_{jt})^{\delta/(1-\sigma)}$, and a conditional choice probability $\hat{s}_{jt} = \frac{\exp\left(\frac{\theta_j - \lambda(t - r_j) + \tau_t + \xi_{jt}}{1-\sigma}\right)}{D_t^{\sigma} + D_t}$. Then $s_{jt} = \pi_{jt} \cdot \hat{s}_{jt}$.

model, where the cross-elasticity of demand is the same across all alternative movies and the outside option. Therefore, the magnitude of σ pins down the relative size of the market-expansion effect.

The unobserved propensity to like movie j in week t is $\xi_{jt} = \omega_{jt} + \varepsilon_{jt}$, where ε_{jt} is an independent and identically distributed measurement error, and ω_{jt} captures an unobserved demand shock that varies across all movies and weeks. If ignored, ω_{jt} can cause endogeneity bias. The instrumental approach in dealing with the endogeneity problem is explained in Sect. 4.2.

4.2 Identification of Demand Parameters

The within-industry market share $\frac{s_{jt}}{1-s_{0t}}$ and theater availability $\ln(m_{jt})$ are both endogenous with respect to the unobserved movie demand shock ω_{jt}. As in Einav (2007) and Moul (2007b), an instrumental variable approach is used to correct for the endogeneity bias. We use the number of rival movies shown in a given week and the market-share weighted average rival weeks-in-release as instrumental variables. These variables capture a movie's competitive environment, which is assumed to be uncorrelated with the movie's own demand shock ω_{jt}.

Movie production budgets—B_{jt}—are also potentially endogenous. Movies with better screenplays may attract both larger audiences and induce greater studio investments. Following Ferreira et al. (2012), we use the total production budgets of a movie's producing studio in the previous year as an instrument variable. A movie studio with a large budget in the previous year may be constrained in its current-year spending. This instrument generates a source of variation in current year movie budgets without being contaminated by movie qualities.

The estimation exploits the panel nature of the data to identify key demand parameters. A key assumption—supported by the data pattern shown in Fig. 1—is that the underlying seasonal pattern is stable over the years. By comparing movie box-office performances in the same calendar week across different years, we can identify both seasonality parameters τ_t and the market-expansion effect parameter σ. For example, if a high quality movie or an increase in the number of movies causes the industry box-office revenue to increase significantly during a year, then the market-expansion effect must be large and σ is close to zero. However, if the increase in industry box-office revenue is small, then the market-expansion effect is small and σ is close to one.

In general, a movie's overall box-office revenue declines over time. This decline can be attributed to the shrinking pool of available consumers who have not already watched a movie. This idea is captured by a decay parameter, λ, in consumer utility. Because the model assumes that all movies follow the same utility decay pattern conditional on observables, parameter λ can be identified by examining a cross-section of movies.

A second possible explanation for the decline in box-office revenue is the decline in theater availability. As theater availability falls, a movie is gradually eliminated from consumers' consideration sets. For a fixed λ, the changes in a movie's market shares and its share of theaters over time jointly identify the theater availability parameter δ.

A potential challenge in estimating δ is that we do not observe the types of theaters. Ticket prices of movies that play in 3-D and IMAX theaters are higher in general. These movies may have inflated market shares in our measure. An implicit assumption is that the 3-D and IMAX theater owners make movie showing decisions in a similar manner to regular theater owners. This assumption means that movie shares in different types of theaters fall in a similar way. Then, the identification of δ is unbiased because different ticket prices do not affect the changes in a movie's market share over time.

4.3 Budget Decisions

On the supply side, movie producers make production budget decisions, taking into account their effects on movie qualities and consequently movie demand. Therefore, biased demand estimates can affect the supply-side predictions. To show the consequences of ignoring theater availability, we model budget decisions similar to that in Ferreira et al. (2012). The profit earned from releasing a movie j is

$$\Pi_j = \left(\sum_{t \in T_j} p_t \cdot M_t \cdot W(t - r_j) s_{jt}(\theta_j(B_j)) \right) \cdot (q_j + R(B_j)) - B_j. \tag{6}$$

In this specification: p_t is the average ticket price in week t; B_j is movie j's production budget; and M_t is the U.S. population size in week t. As defined in Eqs. (2) and (4), movie market share s_{jt} is a function of movie quality θ_{jt}, which in turn is a function of production budget B_j.

Movie producers and theater owners typically arrange revenue-sharing contracts. The movie producer's share is $W(t - r_j)$, which is a function of the movie's number of weeks-in-release. Typically, as is reported in Filson et al. (2005), the movie studio's box-office revenue share is larger (70–90 %) in the opening weeks, then it gradually falls over the weeks-in-release. At the end of a movie's theatrical run, the movie producers can receive a share that is as low as 30%.

Furthermore, studios make budget decisions that take into account other sources of producer income, such as DVD sales, streaming, and merchandising. We assume that a movie's revenues that are earned outside of theaters are proportional to the studio's theater earnings. The factor q_j captures this additional earning ratio.

In addition to the domestic market, a movie can generate box-office returns in international markets. The data provide the ratio of world-wide to the U.S. domestic box-office revenues: R_j.[13] The ratio can be correlated with the production budget, because a movie studio is more likely to promote a blockbuster movie overseas

[13] The ratio R_j is at least 1. It equals 1 when a movie has no international sales.

when it has a large production budget. We use the following regression to capture this relationship:

$$R_j = \alpha_R + \beta_R \ln(B_j) + \tilde{\mu}_g + \tilde{\psi}_y + \tilde{\epsilon}_j, \tag{7}$$

where β_R captures the relationship between budget B_j and the international sales ratio R_j. This regression also controls for movie genre and release year fixed effects, which are represented by $\tilde{\mu}_g$ and $\tilde{\psi}_y$ respectively. Similar to the demand estimation, we use the movie studio's total production budgets in the previous year as an instrumental variable to correct for potential endogeneity bias.

A movie producer chooses B_j to maximize profit as specified in Eq. (6). We assume that movie producers have perfect foresight during the entire length of a movie's theatrical run. They can perfectly predict the number and qualities of rival movies, the number of theaters in which a movie will be exhibited, and the demand shocks in each week-in-release.[14]

Based on the first order condition for profit maximization, movie j's production budget $\hat{B}_j$ is

$$\begin{aligned} \hat{B}_j =& \beta \cdot (R_j + q_j) \sum_t p_t M_t W(t - r_j) \frac{s_{jt}}{1-\sigma} \left[1 - s_{jt} \left(\sigma \frac{s_{0t}}{1 - s_{0t}} + 1 \right) \right] \\ &+ \beta_R \sum_t p_t M_t W(t - r_j) s_{jt}. \end{aligned} \tag{8}$$

The derivation of Eq. (8) is detailed in "Appendix 2".

5 Results

This section presents the estimation results, and discusses the implications of the findings.

5.1 Estimation

Table 6 presents the estimation results of two nested-logit models, based on Eq. (5). Model 2 is the model specified in Eq. (5). Model 1 is the same as Model 2, except that Model 1 does not control for theater availability: Parameter δ is set to zero in Model 1.

The sample period is from late-1994 to mid-2017. For some movies in 1994 and 2017, the data do not contain all weekly observations during their theatrical runs. This data censoring can potentially affect the estimation of movie quality parameters, such as β. Also, we do not have the aggregate studio budgets in 1994. For these

[14] Removing this assumption would add more noise to the prediction of movie production budgets for individual movies. Because we only consider the average industry budget spending in the ensuing analysis, more noise at the individual level is less of a concern.

Table 6 Estimation result

Parameter	Description	Model 1 (without availability)	Model 2 (with availability)
β	Production budget	0.363***	0.184***
		(0.077)	(0.062)
δ	Theater availability		0.447***
			(0.064)
λ	Utility decay	0.243***	0.195***
		(0.049)	(0.034)
σ	Market-expansion effect	0.509***	0.373***
		(0.082)	(0.024)
μ_2	Genre: Comedy	0.073	− 0.023
		(0.055)	(0.044)
μ_3	Genre: Drama	0.119*	0.098*
		(0.068)	(0.053)
μ_4	Genre: Horror/Suspense	− 0.001	−0.087**
		(0.058)	(0.046)
μ_5	Genre: Others	0.069	− 0.013
		(0.105)	(0.085)
α	Constant	− 4.414***	−3.636***
		(0.409)	(0.318)
R^2		0.911	0.943

All of the above estimations control for week specific fixed effects. Standard errors, corrected for heteroskedasticity and arbitrary auto-correlation within a movie, are given in parentheses. Three (***), two (**), and one (*) stars indicate statistical significance at the 1%, 5%, and 10% level respectively

reasons, observations of those movies that are released in or before 1995 and those released in 2017 are not included in the estimation. However, they are considered in the instrumental variables. In particular: Their production budgets are used to calculate the aggregate studio budgets; they are counted in the number of rival movies; and their weeks-in-release are used to calculate the average age of competitors.

Bootstrapping is used for inference that is robust to heteroskedasticity. Standard errors are clustered at the individual movie level to allow for arbitrary auto-correlation within the cluster.

5.2 Result Implications

Theater availability is important because staying in theaters longer allows a movie to remain in consideration sets, thus boosting its box-office performance (holding everything else constant). In Model 2, the estimated theater availability parameter δ is positive and significant at the 1% level. Ignoring theater availability, Model 1 has an estimated production budget parameter, β, that is almost twice as high as that in

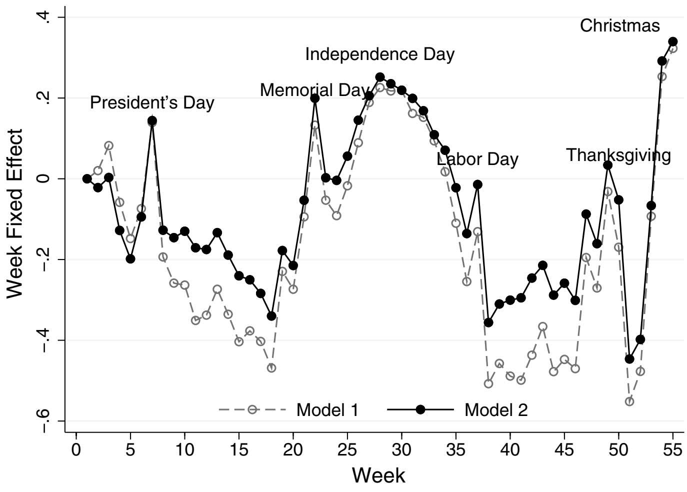

Fig. 2 Estimated seasonality

Model 2. Thus, Model 1 mis-attributes some of a movie's theater staying power to its perceived quality (as represented by its budget).

The estimated utility decay parameter λ is also biased in Model 1. Using the theater share as a proxy, Model 2 can capture the idiosyncratic downward trend in box-office revenue specific to a movie. The average rate of decay in utility, λ, is therefore smaller in Model 2 than in Model 1.

In addition, correcting for changes in market share that are due to theater availability, Model 2 predicts that within-industry competition has a relatively smaller effect on a movie's demand. As a result, the estimated parameter σ is closer to zero in Model 2 than in Model 1, which indicates that the market-expansion effect is larger in Model 2.

Both models include week fixed effects. Figure 2 compares the estimated seasonality, which is very similar across the two models. This shows that theater availability does not much affect the general seasonal patterns in the industry.

Both models also include movie genre fixed effects μ_g. In Table 6, the omitted baseline is the Action/Adventure genre. Movies in the Drama genre are predicted to have significantly higher average quality than the Action/Adventure movies in both models. The average quality of the Horror/Suspense movies is slightly lower than the baseline in Model 2. The average qualities of all other genres are not significantly different from the baseline in either model.

We include release-year fixed effects to control for aggregate industry changes. Figure 3 shows that the relative market share of an average movie to outside options has fallen over the sample years in both models. The fall is especially precipitous from 1996 to 2005, which indicates that it is likely due to home media's becoming

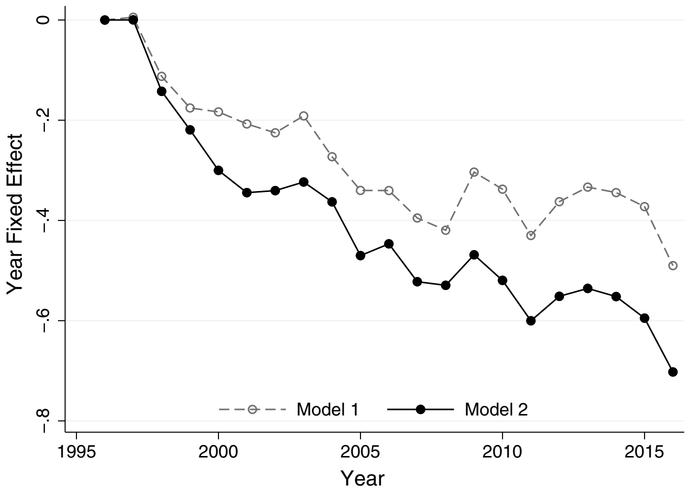

Fig. 3 Estimated release-year fixed effect

more popular in the sample period.[15] Meanwhile, other aggregate factors—such as the increasing prevalence of online movie piracy, the decrease in consumer incomes due to the burst of the dot-com bubble, and the increase in average movie ticket prices—could also have contributed to the downward trends in Fig. 3.

As was discussed previously, we use instrumental variables that reflect to competitive environments and previous-year studio budgets to correct for endogeneity biases. The estimated coefficients have the expected signs in the first-stage regressions with respect to the instrumented variables. For example, a movie's within-industry market share is negatively correlated with the number of competitors and positively correlated with the average weeks-in-release of rival movies. In addition, F-tests of excluded instruments are conducted, and the weak instrumental variables hypotheses are rejected at the 1% level.

5.2.1 Robustness Tests

To test the validity of assuming linear decay in utility, a quadratic term $-\lambda_2(t - r_j)^2$ is added to Model 2. Estimating this alternative model yields $\lambda = 0.410$ and $\lambda_2 = -0.019$. This means that the movie utility decay is faster in the beginning, then slows down towards the end of theatrical runs. The corresponding coefficient on theater availability is $\delta = 0.686$. This suggests that the importance of theater availability is underestimated when we assume linear decay in utility. When predicting production budgets, we use the more conservative estimate of δ presented in Table 6.

[15] For example, five major Hollywood studios formed a joint movies-on-demand venture, which allowed streaming via broadband Internet for the first time in 2002. In the same year, Netflix, reaching 1 million subscribers, made its initial public offering.

The denominator of our proxy for theater availability—the total number of theaters—can reflect some structural changes in the movie exhibition industry and potentially confound the estimation. As a robustness check, the logarithm of the number of estimated screens is instead used to proxy for theater availability when estimating Model 2.[16] The resulting estimates do not change very much. For example, the theater availability parameter δ changes from 0.3730 to 0.3747.

To make sure that no structural changes in consumer demand occurred in the sample period, we re-run Model 2 with the use of different sub-samples by time periods. The differences in the relevant parameters for consumer utility decay (λ), market-expansion effect (σ), and theater availability (δ) are not significantly different between the subsample after year 2000 and the pooled sample.[17] We also re-estimate Model 2 with the use of different sub-samples by genres. The estimates of λ, σ, and δ are not significantly different between the subsample containing only the Action/Adventure movies and the pooled sample.[18]

5.3 Studio Budget Predictions

To understand the importance of theater availability in making supply-side predictions, we compare the predicted production budgets that are based on Eq. (8) with the actual budgets in the data.

In constructing the predicted production budget, we take p_t, M_t, s_{jt}, s_{0t}, and R_j from the data, and use β and σ from the demand estimates. We assume that the movie producers' share of box-office revenue—$W(t - r_j)$—is linear in movie j's weeks-in-release $(t - r_j)$. Following Filson et al. (2005), we set the share W to be 70% in the opening week, and assume it decreases by 5% every week with a floor of 30%. In Eq. (8), the coefficient β_R captures the relationship between production budget B_j and international box-office ratio R_j. Using Eq. (7), β_R is estimated to be 0.109 with a standard error of 0.019. The full regression results are in "Appendix 3".

According to Einav (2007), only 15–35% of the studios' total revenues come from domestic box-office sales. Our data do not record other sources of producer income, such as DVD sales, streaming, and merchandising. Therefore, to make predictions on production budget investment, we assume that domestic box-office sales accounts for at most 25% of the studios' total revenues. Therefore, in Eq. (8), studio incomes from other sources are three times of the domestic box-office revenues, or $q_j = 3$.

[16] Since the data do not report a movie's weekly number of screens, we approximate it by using the product of its weekly number of theaters and the average number of screens per theater as reported in Table 3.

[17] The difference in parameter λ between the subsample after the year 2000 and the pooled sample is – 0.005 (0.174); the difference in σ is 0.045 (0.178); and the difference in δ is −0.053 (0.895). Using a Wald test, these differences are not significantly different from zero: $P > \chi^2 = 0.501$.

[18] The difference in λ between the subsample that contains only the Action/Adventure movies and the pooled sample is 0.009 (0.042); the difference in σ is 0.008 (0.045); and the difference in δ is 0.020 (0.195). Using a Wald test, these differences are not significantly different from zero: $P > \chi^2 = 0.809$.

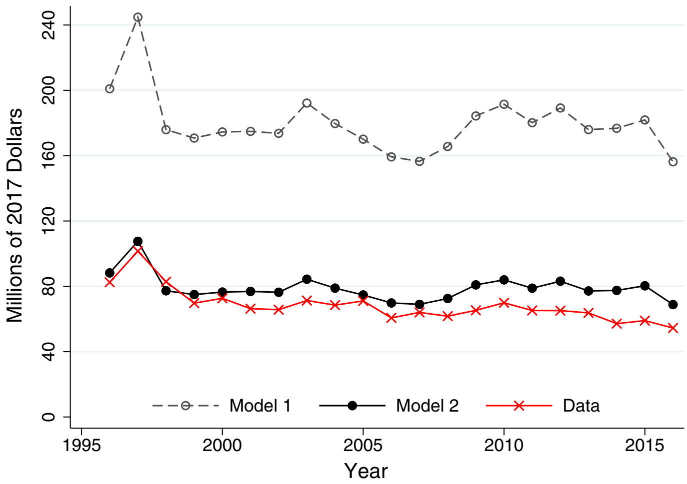

Fig. 4 Average production budget by movie release year

Equation (8) is used to predict the production budget of every movie in the data. We then calculate the average movie production budget in year *y* by averaging the predicted production budgets of all the movies released in year *y*. Similarly, we calculate the mean of actual budgets of all the movies released in year *y* from the data.

Figure 4 compares the average movie production budgets that are predicted by the models to those constructed from the data. The solid line with cross markers represents the average budgets in the data, the dashed line with hollow markers represents the Model 1 predictions, and the solid line with round markers represents the Model 2 predictions.

Models 1 and 2 have different predictions because the different demand estimates from the two models lead to different predicted marginal returns. Table 6 shows that parameter β is higher in Model 1, which translates into a higher marginal return to budget investment in Model 1. Therefore, Model 1, without theater availability, consistently predicts higher average production budgets than does Model 2.

Figure 4 suggests that Model 2 predictions match reasonably well with the actual average production budget in the data. Therefore, our findings imply that theater availability is an important consideration when movie studios make production budget decisions.

We also quantify the magnitude of projected over-investment in production budgets if theater availability is ignored. Compared to the actual budgets, the projected average over-investment in a movie is \$111.25 million, which represents over 60% of the actual production budget.

6 Conclusion

In this paper, we first estimate the impact of theater availability on box-office revenue with the use of a structural model of movie demand, and then highlight the importance of incorporating theater availability in movie studios' production budget decisions. We find that the model that takes into account theater availability is more consistent with the data, and the model that ignores theater availability would over-predict the production budgets.

Provided with suitable data, the next logical step is to extend our research to understand the impact of movie exhibition industry consolidation on theater availability. Interestingly, large chain theaters—such as AMC, Carmike, and Regal—have expanded tremendously in recent decades, which has contributed to a trend of consolidation in the movie exhibition industry. The largest five U.S. exhibitors now own more than half of all screens in the U.S. market. These large operators tend to increase the scale of operation and to upgrade the quality of their theaters. Large movie exhibitors may use their market power to force film distributors to grant them more favorable terms, thereby potentially blocking smaller competitors from screening certain movie titles.[19] Our future research can contribute to the literature of movie theater competitions, and complement the existing papers such as Davis (2006a, b).

Acknowledgements We thank the editor and two anonymous referees for making our paper much better. We thank the participants for their useful comments in IIOC. The usual disclaimer applies.

Appendix 1

In this appendix, we provide the necessary steps in deriving Eqs. (4) and (5) in the main text. Let $v_{jt} = \theta_j - \lambda(t - r_j) + \xi_{jt} + \delta \ln(m_{kt})$, where m_{jt} is the proxy for theater availability, and θ_j is movie j's perceived quality specified in Eq. (2) of the main text. Consumer i's taste shock in week t, $\varsigma_{it} + (1-\sigma)\varepsilon_{ijt}$, follows an extreme value distribution. In addition, the outside option is always considered by consumers, so $v_{0t} = -\tau_t$. Using the standard nested logit specification (McFadden 1978), the probability of consumers choosing movie j in week t is:

$$s_{jt} = \frac{\exp(v_{jt}/(1-\sigma)) \cdot [\sum_{k \in J_t} \cdot \exp(v_{kt}/(1-\sigma))]^{-\sigma}}{[\exp(-\tau_t/(1-\sigma))]^{(1-\sigma)} + [\sum_{k \in J_t} \exp(v_{kt}/(1-\sigma))]^{(1-\sigma)}}.$$

[19] Recently, multiple lawsuits have been filed against the large theater chains; the suits allege that these companies use their market power to coerce film distributors to grant them "clearances." Clearance means that theaters can request that the film distributors decline to license their films to theaters that are located in zones that the large chains deem competitive, because of geographic proximity or shared audiences. See reports from the Washington Post.

In this specification, the outside option is in a different nest and can have a different substitutability from all the movies. The fraction of population that would go to theaters in given week is:

$$1 - s_{0t} = \sum_{k\in J_t} s_{kt} = \frac{\sum_{k\in J_t} \exp(v_{kt}/(1-\sigma)) \cdot [\sum_{k\in J_t} \cdot \exp(v_{kt}/(1-\sigma))]^{-\sigma}}{[\exp(-\tau_t/(1-\sigma))]^{(1-\sigma)} + [\sum_{k\in J_t} \exp(v_{kt}/(1-\sigma))]^{(1-\sigma)}}$$
$$= \frac{[\sum_{k\in J_t} \exp(v_{kt}/(1-\sigma))]^{(1-\sigma)}}{[\exp(-\tau_t/(1-\sigma))]^{(1-\sigma)} + [\sum_{k\in J_t} \exp(v_{kt}/(1-\sigma))]^{(1-\sigma)}}.$$

Correspondingly, the proportion of population choosing the outside option is:

$$s_{0t} = \frac{[\exp(-\tau_t/(1-\sigma))]^{(1-\sigma)}}{[\exp(-\tau_t/(1-\sigma))]^{(1-\sigma)} + [\sum_{k\in J_t} \exp(v_{kt}/(1-\sigma))]^{(1-\sigma)}}$$
$$= \frac{[\exp(\tau_t/(1-\sigma))]^{(\sigma-1)} \cdot [\sum_{k\in J_t} \exp(v_{kt}/(1-\sigma))]^{\sigma}}{[\exp(\tau_t/(1-\sigma))]^{(\sigma-1)} \cdot [\sum_{k\in J_t} \exp(v_{kt}/(1-\sigma))]^{\sigma} + \sum_{k\in J_t} \exp(v_{kt}/(1-\sigma))}$$
$$= \frac{[\exp(\tau_t/(1-\sigma))]^{\sigma}[\sum_{k\in J_t} \exp(v_{kt}/(1-\sigma))]^{\sigma}}{[\exp(\tau_t/(1-\sigma))]^{\sigma}[\sum_{k\in J_t} \exp(v_{kt}/(1-\sigma))]^{\sigma} + \exp(\tau_t/(1-\sigma)) \cdot [\sum_{k\in J_t} \exp(v_{kt}/(1-\sigma))]}$$
$$= \frac{D_t^{\sigma}}{D_t^{\sigma} + D_t},$$

where

$$D_t = \exp(\tau_t/(1-\sigma)) \cdot \sum_{k\in J_t} \exp(v_{kt}/(1-\sigma)) = \sum_{k\in J_t} \exp((v_{kt} + \tau_t)/(1-\sigma)).$$

The within-industry market share of movie j in week t is

$$\frac{s_{jt}}{1 - s_{0t}} = \frac{\exp(v_{jt}/(1-\sigma)) \cdot [\sum_{k\in J_t} \cdot \exp(v_{kt}/(1-\sigma))]^{-\sigma}}{[\sum_{k\in J_t} \exp(v_{kt}/(1-\sigma))]^{(1-\sigma)}}$$
$$= \frac{\exp(v_{jt}/(1-\sigma))}{\sum_{k\in J_t} \exp(v_{kt}/(1-\sigma))}$$
$$= \frac{\exp(\tau_t/(1-\sigma)) \exp(v_{jt}/(1-\sigma))}{\exp(\tau_t/(1-\sigma)) \sum_{k\in J_t} \exp((v_{kt})/(1-\sigma))}$$
$$= \frac{\exp((v_{jt} + \tau_t)/(1-\sigma))}{D_t}.$$

Furthermore,

$$\frac{s_{jt}}{s_{0t}} = \frac{\exp(v_{jt}/(1-\sigma)) \cdot [\sum_{k \in J_t} \cdot \exp(v_{kt}/(1-\sigma))]^{-\sigma}}{[\exp(-\tau_t/(1-\sigma))]^{(1-\sigma)}}$$
$$= \frac{\exp(v_{jt}/(1-\sigma)) \cdot \exp(\tau_t/(1-\sigma))}{[\exp(\tau_t/(1-\sigma))]^{\sigma}[\sum_{k \in J_t} \cdot \exp(v_{kt}/(1-\sigma))]^{\sigma}}$$
$$= \frac{\exp((v_{jt}+\tau_t)/(1-\sigma))}{D_t^{\sigma}}$$
$$= \frac{\exp((v_{jt}+\tau_t)/(1-\sigma))}{(\exp((v_{jt}+\tau_t)/(1-\sigma)))^{\sigma}} \left(\frac{\exp((v_{jt}+\tau_t)/(1-\sigma))}{D_t} \right)^{\sigma}$$
$$= (\exp((v_{jt}+\tau_t)/(1-\sigma)))^{1-\sigma} \cdot \left(\frac{s_{jt}}{1-s_{0t}} \right)^{\sigma}.$$

Taking a log-transformation, we have

$$\ln(s_{jt}) - \ln(s_{0t}) = (1-\sigma)\ln\left(\exp((v_{jt}+\tau_t)/(1-\sigma))\right) \cdot \sigma \ln\left(\frac{s_{jt}}{1-s_{0t}}\right)$$
$$= \delta \ln(m_{jt}) + \theta_j - \lambda(t-r_j) + \tau_t + \sigma \ln\left(\frac{s_{jt}}{1-s_{0t}}\right) + \xi_{jt}$$
$$= \alpha + \beta \ln(B_j) + \mu_g + \psi_y - \lambda(t-r_j) + \tau_t + \delta \ln(m_{jt})$$
$$+ \sigma \ln\left(\frac{s_{jt}}{1-s_{0t}}\right) + \epsilon_j + \xi_{jt}.$$

Also from the derivation of s_{0t}, we have $1-s_{0t} = \frac{D_t}{D_t^{\sigma}+D_t}$. It follows that

$$s_{jt} = \frac{s_{jt}}{1-s_{0t}} \cdot (1-s_{0t}) = \frac{\exp((v_{jt}+\tau_t)/(1-\sigma))}{D_t} \cdot \frac{D_t}{D_t^{\sigma}+D_t}$$
$$= \frac{\exp\left(\frac{\theta_j - \lambda(t-r_j)+\tau_t+\delta\cdot\ln(m_{jt})+\xi_{jt}}{1-\sigma}\right)}{D_t^{\sigma}+D_t}$$

Appendix 2

In this appendix, we show the detailed derivation of Eq. (8) in the main text. Let $\chi_{jt} = \exp\left(\frac{\theta_j - \lambda(t-r_j)+\tau_t+\delta\cdot\ln(m_{jt})+\xi_{jt}}{1-\sigma}\right)$, then from the "Appendix 1", we know that $s_{jt} = \chi_{jt}/(D_t^{\sigma}+D_t)$ and $\frac{d\chi_{jt}}{d\theta_j} = \frac{\chi_{jt}}{1-\sigma}$. In addition, we know that $D_t = \sum_k \chi_{kt}$, where $\frac{\partial D_t}{\partial \chi_{jt}} = 1$.

Therefore,

$$\begin{aligned}\frac{ds_{jt}}{d\theta_j} =& \frac{d\chi_{jt}}{d\theta_j}(D_t^\sigma + D_t)^{-1} - \chi_{jt}(D_t^\sigma + D_t)^{-2}(\sigma D_t^{\sigma-1} + 1)\frac{d\chi_{jt}}{d\theta_j} \\ =& \frac{d\chi_{jt}}{d\theta_j}(D_t^\sigma + D_t)^{-1}\left[1 - \chi_{jt}(D_t^\sigma + D_t)^{-1}(\sigma D_t^{\sigma-1} + 1)\right] \\ =& \frac{s_{jt}}{1-\sigma}[1 - s_{jt}(\sigma D_t^{\sigma-1} + 1)].\end{aligned}$$

Notice that $s_{0t} = \frac{D_t^\sigma}{D_t^\sigma + D_t}$, this means that $D_t = \left(\frac{s_{0t}}{1-s_{0t}}\right)^{\frac{1}{\sigma-1}}$, so the above becomes

$$\frac{ds_{jt}}{d\theta_j} = \frac{s_{jt}}{1-\sigma}\left[1 - s_{jt}\left(\sigma\frac{s_{0t}}{1-s_{0t}} + 1\right)\right].$$

From Eq. (2) of the main text, we also know that $\frac{\partial\theta_{jt}}{\partial B_j} = \beta\frac{1}{B_j}$.

Assuming studio profit $(R(B_j) + q_j)\sum_t p_{jt}M_tW(t - r_j)s_{jt} - B_j$, where $R(B_j)$ is the worldwide to domestic box-office ratio and q_j captures the additional revenue source outside of theaters. Then the first order condition is

$$\begin{aligned}&\beta\frac{1}{B_j}\cdot(R(B_j) + q_j)\sum_t p_{jt}M_tW(t - r_j)\frac{s_{jt}}{1-\sigma}\left[1 - s_{jt}\left(\sigma\frac{s_{0t}}{1-s_{0t}} + 1\right)\right] \\ &+ \beta_R\frac{1}{B_j}\sum_t p_{jt}M_tW(t - r_j)s_{jt} = 1.\end{aligned}$$

We rearrange the above to get Eq. (8) in the manuscript.

Appendix 3

The international box-office ratio is a function of production budget B_j, a movie's genre, and release year. We specify the regression model as the follows:

$$R_j = \alpha_R + \beta_R\ln(B_j) + \tilde{\mu}_g + \tilde{\psi}_y + \tilde{\epsilon}_j$$

To control for the potential endogeneity, we use the total production budgets of the movie producer in the previous year as a instrument variable. The regression results are presented in the Table 7.

Table 7 International box-office ratio regression result

Parameter	Description	Estimated coefficients
β_R	Production budget	0.110*** (0.019)
$\tilde{\mu}_2$	Genre: Comedy	− 0.651*** (0.024)
$\tilde{\mu}_3$	Genre: Drama	− 0.522*** (0.025)
$\tilde{\mu}_4$	Genre: Horror/Suspense	− 0.390*** (0.027)
$\tilde{\mu}_5$	Genre: Others	− 0.568*** (0.044)
$\tilde{\psi}_{1997}$	Release Year: 1997	− 0.114** (0.053)
$\tilde{\psi}_{1998}$	Release Year: 1998	− 0.204*** (0.048)
$\tilde{\psi}_{1999}$	Release Year: 1999	− 0.218*** (0.046)
$\tilde{\psi}_{2000}$	Release Year: 2000	− 0.139*** (0.046)
$\tilde{\psi}_{2001}$	Release Year: 2001	− 0.139*** (0.046)
$\tilde{\psi}_{2002}$	Release Year: 2002	− 0.127*** (0.046)
$\tilde{\psi}_{2003}$	Release Year: 2003	0.031 (0.046)
$\tilde{\psi}_{2004}$	Release Year: 2004	0.072 (0.046)
$\tilde{\psi}_{2005}$	Release Year: 2005	0.168*** (0.046)
$\tilde{\psi}_{2006}$	Release Year: 2006	0.251*** (0.046)
$\tilde{\psi}_{2007}$	Release Year: 2007	0.348*** (0.047)
$\tilde{\psi}_{2008}$	Release Year: 2008	0.409*** (0.046)
$\tilde{\psi}_{2009}$	Release Year: 2009	0.231*** (0.046)
$\tilde{\psi}_{2010}$	Release Year: 2010	0.436*** (0.046)
$\tilde{\psi}_{2011}$	Release Year: 2011	0.540*** (0.046)
$\tilde{\psi}_{2012}$	Release Year: 2012	0.608*** (0.046)
$\tilde{\psi}_{2013}$	Release Year: 2013	0.573*** (0.046)

Table 7 (continued)

Parameter	Description	Estimated coefficients
$\tilde{\psi}_{2014}$	Release Year: 2014	0.523***
		(0.047)
$\tilde{\psi}_{2015}$	Release Year: 2015	0.617***
		(0.048)
$\tilde{\psi}_{2016}$	Release Year: 2016	0.556***
		(0.047)
α_R	Constant	1.723***
		(0.102)

Standard errors are given in parentheses. Three (***), two (**), and one (*) stars indicate statistical significance at the 1%, 5%, and 10% respectively. The omitted genre is Action/Adventure, and the omitted release year is 1996

References

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, *25*(2), 242–262.

Bronnenberg, B. J., & Vanhonacker, W. R. (1996). Limited choice sets, local price response and implied measures of price competition. *Journal of Marketing Research*, *1*, 163–173.

Dalton, J. T., & Leung, T. (2017). Strategic decision-making in Hollywood release gaps. *Journal of International Economics*, *105*, 10–21.

Davis, P. (2006a). Measuring the business stealing, cannibalization and market expansion effects of entry in the U.S. motion picture exhibition market. *The Journal of Industrial Economics*, *54*(3), 293–321.

Davis, P. (2006b). Spatial competition in retail markets: Movie theaters. *The RAND Journal of Economics*, *37*(4), 964–982.

De Vany, A., & Walls, D. (1996). Bose–Einstein dynamics and adaptive contracting in the motion picture industry. *Economic Journal*, *106*(439), 1493–1514.

De Vany, A., & Walls, D. (1997). The market for motion pictures: Rank, revenue and survival. *Economic Inquiry*, *35*(4), 783–797.

Einav, L. (2007). Seasonality in the U.S. motion picture industry. *The RAND Journal of Economics*, *38*(1), 127–145.

Ferreira, F., Petrin, A., & Waldfogel, J. (2012). Trade, endogenous quality, and welfare in motion pictures. Working paper.

Filson, D., Switzer, D., & Besocke, P. (2005). At the movies: The economics of exhibition contracts. *Economic Inquiry*, *43*(2), 354–369.

Fotheringham, A. S. (1988). Note—Consumer store choice and choice set definition. *Marketing Science*, *7*(3), 299–310.

McFadden, D. (1978). Modeling the Choice of Residential Location. Transportation Research Board Research Record No.673.

Moretti, E. (2011). Social learning and peer effects in consumption: Evidence from movie sales. *The Review of Economic Studies*, *78*, 356–393.

Mortimer, J. H. (2008). Vertical contracts in the video rental industry. *The Review of Economic Studies*, *75*(1), 165–199.

Moul, C. (2007a). Measuring word-of-mouths impact on theatrical movie admissions. *Journal of Economics and Management Strategy*, *16*(2), 859–92.

Moul, C. (2007b). Retailer entry conditions and wholesaler conduct: The theatrical distribution of motion pictures. *International Journal of Industrial Organization*, *26*, 966–983.

Wu, J., & Rangaswamy, A. (2003). A fuzzy set model of search and consideration with an application to an online market. *Marketing Science*, *22*(3), 411–434.