# Learning about movies: the impact of movie release types on the nationwide box office

**Xinlei Chen · Yuxin Chen · Charles B. Weinberg**

Received: 24 July 2011 / Accepted: 16 October 2012 / Published online: 29 October 2012


**Abstract** Major Hollywood studios typically release new movies in North America in one of the two ways, wide release or platform release. In this paper, we investigate how release form affects the demand of a new movie after it is nationally released. In particular, we focus on movies for which the platform release is pre-planned to make the problem tractable. We estimate our model using a sample of Hollywood movies that eventually received nationwide release from 1999 to 2003. Our results show that platform release shifts consumers' perception of unobservable movie appeal through its first stage performance, which turns out to be a stronger effect than that of advertising. Meanwhile, we find that the demand for platform movies decays faster than for wide release ones after their national release. Using counterfactual analysis, we find that more than half of the platform movies which later went to national release would have earned higher profits if they had been given a wide release.

**Keywords** Entertainment marketing · Distribution · Word of mouth · Movies



The authors are listed alphabetically and contributed equally to the paper.

X. Chen (✉) · C. B. Weinberg
Sauder School of Business, University of British Columbia, Vancouver, BC, Canada
e-mail: jack.chen@sauder.ubc.ca

C. B. Weinberg
e-mail: charles.weinberg@sauder.ubc.ca

Y. Chen
Kellogg School of Management, Northwestern University, Evanston, IL, USA
e-mail: yuxin-chen@kellogg.northwestern.edu

Hollywood studios release 400–600 movies annually (MPAA 2007) in North America, but only half of them reach the national market.[1] These movies earned (from 1999 to 2003) 97.5 % of the total North American box office revenue. A new movie can reach the national market in one of two ways–wide release or platform release. A "wide release" movie opens nationwide in a large number of theatres, ranging from several hundred to more than three thousand; after that, the number of theatres typically declines over time. In contrast, a platform release movie is first released in a limited number of theatres in selected major cities. Then, either by a pre-planned strategy or in response to box office success while in limited release, the studio gives the movie a national release. This starts the second stage (hereafter referred to as the national release stage) and the movie's subsequent diffusion pattern resembles that of a wide release movie.

Under platform release, a movie can reach the national market through two very different decision processes. In the first case, a studio is targeting the national market but expects that the platform release can generate positive buzz about the movie, therefore improving its box office performance in the national market. In this case the timing of going national is pre-determined, as it is set before the release of the movie. In the second case, the potential appeal is highly uncertain or the studio does not have the funding for a national release. In this case, a successful platform release serves to justify the funding of a marketing campaign and a national release. Although the popular perception is that most platform movies that achieve national release are unexpected "sleepers," this is not the case. In 1999–2003, 121 movies went national after a platform release. Using our definition (see below) of pre-planned national release, 60 % of the platform movies that went national were pre-planned for national release, accounting for 77 % of total box office revenue for those 121 movies.

The central questions of this paper are how a platform release affects demand for the subsequent national release and whether studios can improve their revenues by going directly to wide release in these cases? However, as we will quickly see, obtaining a general answer to this question is difficult. This is because when a national release is not pre-planned, the decision to go to the national market becomes endogenous. In fact, most platform release movies never achieve national release. By only looking at platform movies that eventually went to the national market, we induce sample selection bias in the estimation. The solution to this problem is to structurally model studios' decisions of bringing a platform movie to national market. This turns out to be a challenge task. Given the above concerns and the data availability, in this paper we will take the first step to address an interesting research stream by focusing on platform movies for which the national release is pre-planned. By doing this, we not only avoid any sampling biases from choosing only "successful" platform release movies, but also focus on movies accounting for the majority of the revenues from platform releases.

[1] We will define the national market later in the paper.

ways. First, compared to wide release, platform movies reveal more information at the first stage, which will ultimately shape consumers' expectations of the movie when it reaches national release. Second, as with other perishable and fashion goods, consumers' desire for a new movie declines over time. A platform movie may suffer from obsolescence because consumers have heard about it during its earlier platform release. Both factors may affect consumers' purchase decisions for a new movie.

To understand how release form affects the demand of a new movie after its national release, we develop a formal learning model to capture consumers' decision to watch a movie. In our model, consumers' purchase decisions depend on a movie's perceived appeal, which is a function of movie features. As a movie is a typical experience product (Eliashberg and Sawhney 1994; Nelson 1974), some of a movie's features are unobservable prior to consumption; hence a movie's appeal consists of both appeal from observable features (hereafter observable appeal) and appeal from unobservable features (hereafter unobservable appeal). Consumers try to learn about the unobservable appeal. Two information sets are available to facilitate consumers' learning: (a) prior belief about the unobservable appeal before the national release of the movie, and (b) word of mouth signals from those who have seen the movie once it is available in the market. A major difference between platform and wide release on the movie demand after national release lies in the way that they affect consumers' learning. Specifically, for platform movies, consumers' prior perceptions before national release can shift based on its first stage performance. In addition, the two release types can differ in terms of advertising effect and decay rate.

We estimate our model using a sample of Hollywood movies that eventually received nationwide release from 1999 to 2003. Our results show that platform release shifts consumers' perception of unobservable appeal upward through its first stage performance. Although advertising has similar effects in terms of shaping consumers' prior belief, the first stage performance turns out to have a much stronger effect. Our results reveal that \$1 million in first stage box office revenues shifts consumers' prior belief of unobservable appeal to the same extent as \$44.26 million in advertising expenditure for dramas, \$15.03 million for comedies, and \$13.56 million for action movies. Although there is no significant difference in advertising effect between release types, we find that platform movies decay faster than wide release ones in the national release stage. Thus, it is not obvious as to which release type leads to higher nationwide box office. Using counterfactual analysis, we find that nearly half of the platform movies that eventually went to national release would have earned higher profits by switching to wide release. This provides important information for studios to evaluate their platform release decisions.

Our research contributes substantively to the growing literature on the motion picture industry. While a number of papers have focused on improving the method to forecast a movie's box office revenue, (Sawhney and Eliashberg 1996; Neelamegham and Chintagunta 1999; Elberse and Eliashberg 2003; Anslie et al. 2005), little is most immediately for entertainment products and fashion goods. Even for package goods, understanding the impact of limited distribution in selected markets on the demand of a new product after its nationwide launch is critical when firms decide whether to rollout nationwide at the beginning or to start by building local success. In some ways, a platform release is a special case of test marketing. Methodologically, our paper constructs a structural learning process, which facilitates the counterfactual analysis (Kadiyali et al. 2001).

The rest of the paper is organized as follows. In Sect. 2, we describe the data used in our study and provide background information on the motion picture industry that is relevant to our research. In Sect. 3, we present the model as well as the estimation method. The findings from our empirical analysis are given in Sect. 4 along with the interpretation and discussion of those findings. Finally, we conclude the paper in Sect. 5 and outline future research directions.

## 2 Data

Our sample consists of all movies that were released in North America (domestic market) from January 1999 to September 2003. For each movie, the data contain three sources of information: (a) weekly box office revenue and weekly number of theatres engaged from Variety.com, (b) advertising expenditure from Competitive Media Report (CMR),[2] and (c) movie characteristics such as production budget, genre,[3] MPAA rating, distributor, critics and viewer rating, and Oscar award nominations and winning records, collected from variety.com, IMDB.com, and rottentomatoes.com.

A movie can be released in one of the two formats: wide release or platform release. Figure 1 illustrates the difference between these two release types. As there is no clear-cut definition distinguishing wide release and platform release movies from each other, we empirically define the two types of release for our sample. As can be seen in Fig. 2, 200 is the median number of theatres for our sample. Consequently we begin by assuming 200 as the minimum number of theatres engaged for a movie to go to national release. While 200 theatres covers only a small fraction of the total theatres available, it represents a situation in which the movie is available in most major metropolitan areas in North America. This is confirmed by our conversations with industry experts. As can be seen in Fig. 2, the maximum number of theatres in which a movie is shown is fairly uniformly distributed above 200. This provides reassurance that 200 theatres is a reasonable criterion for national release. Given such a criterion, we define a movie to be a wide release movie if it played in more than 200 theatres in its opening week. Otherwise,

[2] Later on, CMR was bought by TNS.

[3] To avoid including genres in which very few movies appeared, we reduce the number of genres to three, that is, drama, comedy and action. For a movie, we collect the genre keywords from IMDB.COM. The movie's genre is defined based on whichever of the above three genres appears first.

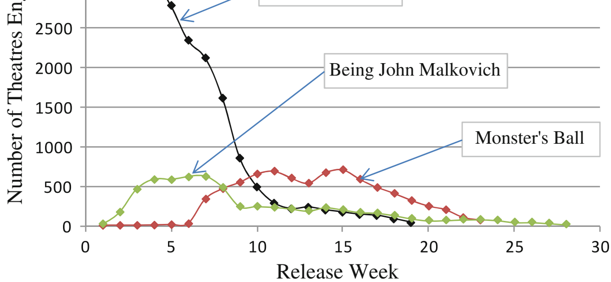


**Fig. 1** Number of theatre engaged over time under different release type

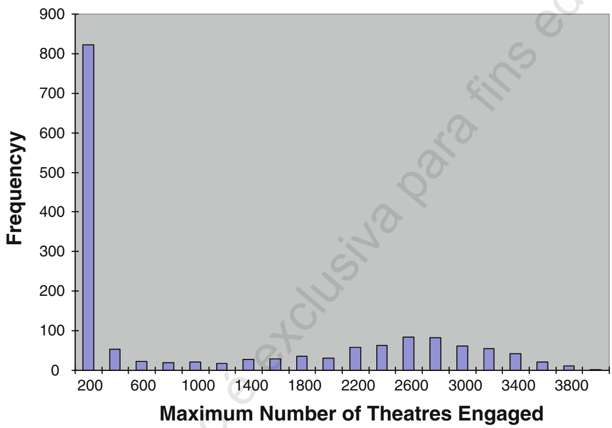


**Fig. 2** Distribution of maximum number of theatres engaged

it is a platform movie. Similarly, for a platform movie, we define the week that it goes to national release as the week when the number of theatres exceeds 200.[4]

As we stated before, there are two scenarios in which a platform movie could go to national release, either pre-planned by the studio or fully depending on its performance in the first stage. In this paper, we only focus on the platform movies which are pre-planned to go national by the studios. Given the set of platform movies which eventually went national, we don't have information on whether the national release is pre-determined or not. However, our conversation with studio managers suggests that if not pre-planned, it would be difficult for studios to switch to national release for a platform movie within a short period of time, given that it typical takes time for studios to negotiate the distribution contract with theatres.



4 We tested our model using both 300 theatres and 400 theatres as a criterion and obtained similar results to those reported in the paper.

From the initial 1,627 movies in our sample, we restrict our analysis to those movies for which we have a complete data set and which finally went to national release pre-determined by the studios, that is, achieved a distribution level of at least 200 theatres within a month of release. The final sample contains 487 movies with 7,122 weekly observations. These movies account for 65.4 % of the total box revenues of movies in our initial sample. Among these movies, 49 are categorized as platform movies which are pre-planned to go national by the studios. Table 1 contains descriptive statistics for both the wide release and platform movies in the sample. As can be readily seen, wide release movies tend to be bigger budget movies with higher advertising expenditures but with shorter life in theatres, on average 3 weeks less than that of platform movies. Wide release movies are more likely to be sequels and to be distributed by a major distributor. In terms of genre, wide release movies are evenly distributed across different categories, while platform movies are dominated by drama and comedy. In terms of winning awards and earning higher ratings, platform movies do better than wide release ones on average. To illustrate the difference between platform movies which are pre-planned to go national and which are not, we also report the descriptive statistics for 21 platform movies which were not pre-planned but eventually went national, that is, they reached 200 theatres after one month. Table 1 shows that there are also significant differences between the two types of platform movies. In general, pre-planned platform movies have a higher budget, and more advertising support and theatre coverage once they go national.

## 3 Model

In this section, we propose a model of movie demand in the national release stage. In discussing our model, we start with the utility specification followed by a description of the process by which consumers learn about movie appeal. Figure 3 illustrates the sequence of events before a consumer decides to watch a movie. Then, we present the empirical specification of the model. Finally, we discuss the identification and estimation method for the model.

### 3.1 Utility specification

At time $t$ after a movie's national release, a consumer decides whether to see the movie or not based on his/her perception of the movie's appeal. As very few people will see a movie more than once in a theatre, we start by assuming that there are no repeat purchases in the movie market. Hence we focus on those customers who have not seen the movie, because once a consumer sees the movie, he/she is out of the market. At a given time $t$, consumer $i$'s utility of going to see a movie $j$ is

[5] We tested our model using platform movies that went to national release within 3 weeks and obtained similar results to those reported in the paper.

| Variable | Mean | SD | Minimum | Maximum | Median |
|---|---|---|---|---|---|
| *Wide release (N = 438)* | | | | | |
| BUDGET(Million $) | 43.09 | 31.44 | 1 | 175 | 35 |
| OSCAR_N | 0.45 | 1.83 | 0 | 16 | 0 |
| OSCAR_W | 0.06 | 0.41 | 0 | 5 | 0 |
| Rating_Critics | 5.13 | 1.29 | 2.2 | 8.2 | 5.1 |
| Rating_Viewer | 5.83 | 1.04 | 2.3 | 8.2 | 5.9 |
| Cumulated box office (Million $) | 55.08 | 55.29 | 1.64 | 403.71 | 35.20 |
| Advertising expenditure (Million $) | 11.89 | 5.18 | 0.01 | 27.08 | 11.86 |
| Sequel | 0.09 | 0.29 | 0 | 1 | 0 |
| Life | 14.19 | 5.96 | 6 | 50 | 13 |
| Major distributor | 0.75 | 0.43 | 0 | 1 | 1 |
| Max number of theatres | 2,396.89 | 717.95 | 214 | 3,876 | 2,525 |
| *Genre* | | | | | |
| Drama | 0.31 | 0.46 | 0 | 1 | 0 |
| Action | 0.31 | 0.46 | 0 | 1 | 0 |
| Comedy | 0.38 | 0.49 | 0 | 1 | 0 |
| *MPAA* | | | | | |
| PG-13 | 0.47 | 0.5 | 0 | 1 | 0 |
| R | 0.4 | 0.49 | 0 | 1 | 0 |
| PG | 0.13 | 0.33 | 0 | 1 | 0 |
| *Pre-planned platform release (N = 49)* | | | | | |
| BUDGET(Million $) | 18.3 | 13.46 | 0.04 | 60 | 15 |
| OSCAR_N | 2.9 | 4.42 | 0 | 19 | 1 |
| OSCAR_W | 0.67 | 1.8 | 0 | 8 | 0 |
| Rating_Critics | 6.99 | 0.8 | 5.2 | 8.4 | 7.1 |
| Rating_Viewer | 7.29 | 0.55 | 6 | 8.6 | 7.3 |
| 1st stage box office (Million $) | 2.42 | 2.06 | 0.15 | 9.38 | 2.08 |
| Cumulated box office (Million $) | 32.61 | 45.43 | 1.59 | 238.59 | 15.51 |
| Advertising expenditure (Million $) | 4.57 | 4.19 | 0.02 | 16.2 | 3.31 |
| Sequel | 0 | 0 | 0 | 0 | 0 |
| Life | 18.51 | 8.49 | 7 | 48 | 17 |
| Major distributors | 0.53 | 0.5 | 0 | 1 | 1 |
| Takeoff | 3.22 | 0.71 | 2 | 4 | 3 |
| Max number of theatres | 997.2 | 683.26 | 224 | 2,701 | 842 |
| *Genre* | | | | | |
| Drama | 0.63 | 0.49 | 0 | 1 | 1 |
| Action | 0.06 | 0.24 | 0 | 1 | 0 |
| Comedy | 0.31 | 0.47 | 0 | 1 | 0 |
| *MPAA* | | | | | |
| PG-13 | 0.27 | 0.45 | 0 | 1 | 0 |
| R | 0.71 | 0.46 | 0 | 1 | 1 |

| Variable | Mean | SD | Minimum | Maximum | Median |
|---|---|---|---|---|---|
| PG | 0.02 | 0.14 | 0 | 1 | 0 |
| *Non pre-planned platform release (N = 21)* | | | | | |
| BUDGET(Million $) | 7.77 | 8.07 | 0.15 | 32 | 5.5 |
| OSCAR_N | 2.57 | 2.54 | 0 | 9 | 2 |
| OSCAR_W | 0.48 | 0.87 | 0 | 2 | 0 |
| Rating_Critics | 7.22 | 0.83 | 4.5 | 8.4 | 7.2 |
| Rating_Viewer | 7.19 | 0.45 | 6.4 | 7.9 | 7.2 |
| 1st stage box office (Million $) | 2.96 | 1.94 | 0.3 | 7.4 | 2.87 |
| Cumulated box office (Million $) | 11.49 | 11.17 | 1.04 | 33.41 | 6.48 |
| Advertising expenditure (Million $) | 1.9 | 2.78 | 0 | 9.42 | 0.73 |
| Sequel | 0 | 0 | 0 | 0 | 0 |
| Life | 16.62 | 6.86 | 8 | 31 | 15 |
| Major distributors | 0.38 | 0.5 | 0 | 1 | 0 |
| Takeoff | 7.81 | 4.79 | 5 | 24 | 6 |
| Max number of theatres | 565.1 | 474.46 | 220 | 2,331 | 493 |
| *Genre* | | | | | |
| Drama | 0.81 | 0.4 | 0 | 1 | 1 |
| Action | 0 | 0 | 0 | 0 | 0 |
| Comedy | 0.19 | 0.4 | 0 | 1 | 0 |
| *MPAA* | | | | | |
| PG-13 | 0.19 | 0.4 | 0 | 1 | 0 |
| R | 0.71 | 0.46 | 0 | 1 | 1 |
| PG | 0.1 | 0.3 | 0 | 1 | 0 |

Budget: Production budget

Oscar_N: Number of Oscar nominations

Oscar_W: Number of Oscar awards

Rating_Critics: Critics rating, 1−10 scale

Rating_Viewer: Viewer rating, 1−10 scale

Sequel: Dummy variable to indicate whether a movie is a sequel

Life: The total number of running weeks after national release

Takeoff: Takeoff week

Major distributor: Dummy, indicating whether a movie is distributed by a major distributor : Paramount, Sony Pictures (Columbia Pictures, TriStar), The Walt Disney Company (Buena Vista, Touchstone, and Hollywood Pictures), Twentieth Century Fox, Universal, and Warner Bros (New Line, Fine Line)

Max Number of Theatres: Maximum number of theatres engaged



$$u_{ijt} = E\left(A^u_{ijt}\right) + A^o_{ijt} + \varepsilon_{ijt} \quad (1)$$

where $E\left(A^u_{ijt}\right)$ is the expectation of a movie's unobservable appeal, $A^o_{ijt}$ is the observable appeal, and $\varepsilon_{ijt}$ is an *i.i.d* random error following the type I extreme value distribution. The basic premise of our model is that consumers are uncertain about a

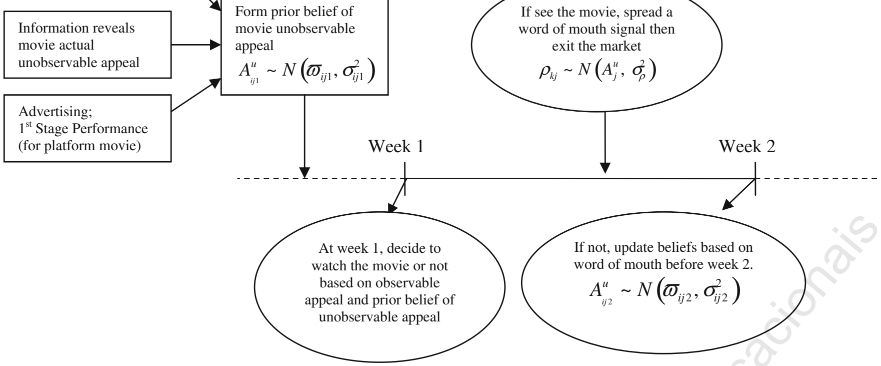


**Fig. 3** Sequence of events prior to watching a nationally released movie

movie's unobservable appeal, and hence attempt to learn about it. $A^u_{ijt}$ is updated each period based on word of mouth from those who have seen the movie, but initially depends on the studio's marketing campaign and the potential viewer's beliefs about movies that are similar (for example, in terms of genre) to this movie. The updating process is described in more detail in the next section. Therefore, $A^u_{ijt}$ is a random variable for consumer $i$. In contrast, $A^o_{ijt}$ is a deterministic term which can be modelled as a function of observable movie features.

Consumer $i$ is going to see movie $j$ if and only if $u_{ijt} \geq u_{i0t}$, where $u_{i0t}$ is the utility from not watching movie $j$ and is normalized to be zero.

## 3.2 Learning process

There are several sources of information that can shape a consumer's belief about the unobservable movie appeal, including his/her prior knowledge/experience with similar movies, the studio's advertising campaign, the first stage performance if the movie has a platform release, and word of mouth spread by movie goers who have seen the movie. We discuss them in detail next.

At a given time $t$, consumer $i$ is uncertain about the unobservable appeal of movie $j$ therefore considers it as a random variable following $A^u_{ijt} \sim N\left(\omega_{ijt}, \sigma^2_{ijt}\right)$, where $\omega_{ijt}$ denotes the expected unobservable appeal of movie $j$ at time $t$, and $\sigma_{ijt}$ captures the uncertainty in consumer $i$'s belief.

### *3.2.1 Evolution of a consumer's belief of unobservable appeal*

At time $t$, a consumer generates his/her belief of unobservable appeal from two information sets. One is his/her prior belief of a movie's unobservable appeal generated from last period, $A^u_{ijt-1} \sim N\left(\omega_{ijt-1}, \sigma^2_{ijt-1}\right)$. The other set of information consists of the signals coming from consumers who have seen the movie, namely friends, write a comment in his/her blog, or rate the movie on the Internet, etc. We generally describe such information spillover as word of mouth from a purchaser. We also assume that viewers spread their word of mouth right after consumption but not in future periods. A person is most likely to want to discuss a movie right after he/she has seen it. Let $n_{j,t-1}$ be the total number of people[6] watching movie $j$ at $t-1$, and $\rho_{kj}$ be the word of mouth signal from each viewer $k$ for movie $j$ with $k = 1\ldots N_{j,t-1}$. Next, we assume word of mouth signals follow a distribution as

$$\rho_{kj} \sim N\left(A_j^u, \sigma_\rho^2\right) \tag{2}$$

where $A_j^u$ is the actual unobservable appeal of movie $j$ and $\sigma_\rho$ is the noise in the word of mouth signals. This is a standard assumption used in the learning literature (Erdem and Keane 1996; Narayanan et al. 2005).

In reality, the number of word of mouth signals received may vary across individuals. However, we do not have information on this. Instead, we assume (a) each individual who watched the movie will spread one signal, and (b) all the $n_{j,t-1}$ signals generated at time $t-1$ will reach an individual who has not yet seen the movie. With such assumptions, the estimate of signal uncertainty in our model captures the average of the total word of mouth effect across individual across time. The mean value of the updated belief is a weighted average of the previous period's mean value of the belief and the current period's word of mouth, where the weighting is based on the inverse value of the variance, that is, a piece of information with high precision (low variance) has more weight in shaping the belief (DeGroot 1970). In other words, consumer $i$ updates his/her prior belief of movie $j$ with word of mouth signals in a Bayesian fashion, that is, $A_{ijt}^u \sim N\left(\omega_{ijt}, \sigma_{ijt}^2\right)$, where

$$\omega_{ijt} = \frac{\frac{\omega_{ijt-1}}{\sigma_{ijt-1}^2} + \frac{1}{\sigma_\rho^2}\sum_{k=1}^{n_{jt-1}} \rho_{kj}}{\frac{1}{\sigma_{ijt-1}^2} + \frac{n_{jt-1}}{\sigma_\rho^2}}, \frac{1}{\sigma_{ijt}^2} = \frac{1}{\sigma_{ijt-1}^2} + \frac{n_{jt-1}}{\sigma_\rho^2} \tag{3}$$

It is straightforward to show that Eq. (3) can be written in the following form:

$$\omega_{ijt} = \frac{\gamma_{ij1}\omega_{ij1} + \sum_{r=1}^{t-1}\sum_{k=1}^{n_{jr}} \rho_{kj}}{\gamma_{ij1} + \sum_{r=1}^{t-1} n_{jr}}, \gamma_{ijt} = \gamma_{ij1} + \sum_{r=1}^{t-1} n_{jr} \tag{4}$$

where $\gamma_{ijt} = \frac{\sigma_\rho^2}{\sigma_{ijt}^2}$, which is the ratio between noise in the word of mouth signal and uncertainty in the belief of unobservable appeal.



[6] Notice that in our data, we do not have the number of people going to see the movie in a certain week. Instead, we have weekly box office revenue for each movie. Given that there is very small variation across time in movie ticket price, we can infer the total number of weekly movie goers through dividing the weekly box office by average ticket price. We have yearly average ticket price from the Motion Picture Association of America (MPAA) from 1999 to 2003. However, we choose not to use these because that will create an artificial change in number of weekly movie goers at the beginning of each year. Instead, we use the average ticket price across 1999 to 2003, which is \$5.59, in our estimation. We tested our model using the yearly average prices and obtained similar results.

We need to specify consumers' prior belief at the opening week of national release. Before the release of a new movie, consumers know very little about the unobservable features of the movie. However, they can generate their belief from two sources of information. Maybe the most important information source comes from the studio's marketing campaign, which includes previews, advertising, and promotion activities such as stars travelling around to promote the movie. For a platform movie, the first stage performance also affects consumers' prior beliefs in the national market. The other source of information is consumers' perception of the unobservable appeal of movies with the similar type, which comes from their past consumption experience.

We set initial uncertainty to be constant across the population and movies, that is, $\sigma_{ij1} = \sigma_1 \, \forall i,j$. We assume that, by the time of national release, the expected unobservable appeal of a movie for a given consumer is a random draw from the population distribution, that is,

$$\omega_{ij1} \sim N\left(B_j + \Delta A_j^u, \sigma_1^2\right) \tag{5}$$

$B_j$, the consumer's intrinsic perception of the unobservable appeal of the movie $j$, is one factor determining the mean of prior belief. And $\Delta A_j^u$ captures how the consumer's belief of the unobservable appeal is shaped by other factors such as advertising and 1st stage performance.

### *3.2.3 Stochastic specification of learning process*

In Eq. (4), a consumer knows the realization of his/her initial belief $\omega_{ij1}$ and word of mouth signals $\rho_{kj}$. As researchers, we have no observation of such information except for their distribution. Therefore, those signals are random variables to us. In order to derive the sample likelihood, we need to write down the stochastic specification of the learning process. From Eq. (2), we can write

$$\rho_{kj} = A_j^u + \sigma_\rho \varsigma_{jt}$$

where $\varsigma_{jt}$ is an *i.i.d.* random error following the standard normal distribution. From this, we can show that

$$\sum_{k=1}^{n_{jt}} \rho_{kj} = n_{jt} A_j^u + \sqrt{n_{jt}\sigma_\rho n_{jt}}$$

where $\eta_{jt}$ is an *i.i.d* random error that follows the standard normal distribution. Similarly, from Eq. (5), we can write down the initial belief as

$$\omega_{ij1} = B_j + \Delta A_j^u + \sigma_1 \xi_{ij1} = B_j + \Delta A_j^u + \frac{\sigma_\rho}{\sqrt{\gamma_{j1}}} \xi_{ij1}$$

where $\xi_{ij1}$ is an *i.i.d* random error that follows the standard normal distribution.

$$\omega_{ij1} = \frac{A_j^u N_{jt-1} + \gamma_{j1}\left(B_j + \Delta A_j^u\right) + \sqrt{\gamma_{j1}}\sigma_1 \xi_{ij1} + \sum_{r=1}^{t-1} \sqrt{n_{jr}}\sigma_\rho n_{jr}}{\gamma_{j1} + N_{jt-1}} \tag{6}$$

where $N_{jt-1} = \sum_{r=1}^{t-1} n_{jr}$.

## 3.3 Empirical specification

### *3.3.1 Observable appeal*

By definition, observable appeal is based on characteristics that are readily available to potential movie goers and can unambiguously be determined. Based on these criteria and the empirical results in the literature, we model a movie's observable appeal as a function of observable movie features in the following equation: $A_{ijt}^0 = \beta_{i1} Genre_j + \beta_2 Z_j + (\beta_{30} Platform_j + \beta_{31} Genre_j) * age_{jt} + \beta_4 * Theatre_{jt} + \beta_5 Season_{jt}$ where $Genre_j$ is a vector of dummy variables to indicate the genre of movie $j$ and $Z_j$ is a vector of dummies including *sequel*, *MPAA* rating, and *major distributor*.[7] Genre has been shown to be a significant factor associated with a movie's appeal in many studies (e.g., Einav 2007). However, as individuals vary in their preference for genre, we allow consumer heterogeneity in their preference for different genres. Specifically, we assume

$$\beta_{i1} \sim N\left(\bar{\beta}_{i1}, \sigma_{\beta_1}^2\right)$$

$Age_{jt}$ is defined as the number of weeks the movie has played at time $t$ in the national market. This term captures the decline of attractiveness of a new movie over time, which is common for new products with short life cycles like movies (see, for example, Krider and Weinberg 1998). The interaction between age and platform captures the possibility that after the first stage, a platform movie may become obsolete faster than a wide release movie. The interaction between age and genre captures the fact that consumers may lose interest faster for some types of movies than for others. $Theatre_{jt}$ is the number of theatres engaged for movie $j$ at time $t$, which captures the extent to which a movie's availability affects a consumer's choice (Sawhney and Eliashberg 1996). A movie's MPAA rating and its being a sequel have been found in a number of studies to be related to a movie's appeal (e.g., Einav 2007 and Basuroy et al. 2006 respectively). As demonstrated in Einav (2007), people find it more appealing to attend a movie when they have more free time, typically during the summer and around holidays. $Season_{jt}$ is a vector of dummy variables including monthly dummies for June, July, and August, which are considered as the high season for movies, and holiday dummies to indicate major



[7] Major Distributor indicates whether a movie is distributed by Paramount, Sony Pictures (Columbia Pictures, TriStar), The Walt Disney Company (Buena Vista, Touchstone, and Hollywood Pictures), Twentieth Century Fox, Universal, or Warner Bros (New Line, Fine Line).

A set of information that is likely observable by movie goers before watching a movie but is not included in our specification of observable appeal is the information on movie ratings from critics and viewers. Instead, we assume that the mean ratings from critics and viewers are correlated with the actual mean of the unobservable appeal as discussed next. The reasons are as follows. First, different potential movie goers may have observed different ratings, and it is well known that there is considerable divergence in the evaluations of a movie. Moreover, people respond in different ways to the set of reviews to which they are exposed. Therefore, the exact rating information accessed by movie goers is ambiguous and its impact is uncertain. Second, critics' and viewers' ratings result from the movie watching experiences of movie critics and other viewers. Such experiences reflect the unobservable appeal of a movie but are not directly observable by potential movie goers. Therefore, we feel it is more reasonable to regard the mean of critics and viewers ratings as indicators of a movie's unobservable appeal instead of having them directly in the specification of the utility from observable movie appeal.

### *3.3.2 Unobservable appeal*

By definition, a movie's unobservable appeal will be reflected by its unobserved features. The unobservable appeal is the part of movie-related utility that is known for certain to consumers only by watching it. Although it is unobservable and therefore difficult to measure directly, the unobservable appeal may be correlated with some indicators such as awards, for example Oscars, and ratings from critics and viewers. Also, the investment in the movie, that is, the production budget can correlate with a movie's unobservable features. Therefore, we model the unobservable appeal as a function of these indicators. It is worth noting that although some of those indicators could be observed by consumers at certain stage of the movie life cycle, consumers are still uncertain of the correlation between these indicators and the unobservable appeal. Specifically, we assume that

$$A_j^u = a_0 + a_1 * Oscar_j + a_2 * Rating + a_3 * Budget_j + a_4 * Budget_j^2$$

where $Oscar_j$ includes total number of Oscar nominations and awards. By including both Oscar nominations and awards, we are able to allow for different levels of relative appeal. Although Oscar nominations and awards are not known at the time the movie is released, they are likely correlated with unobservable appeal. Hence, they are included in the equation for unobservable appeal. $Rating_j$ includes ratings from critics and viewers respectively,[8] and finally $Budget_j$ is the production budget of the movie. We included a quadratic term for the production budget to allow for the possibility of nonlinear effects.

[8] We only have the average critics' and viewers' ratings in our data. If we had information on the distribution of critics' and viewers' ratings, we could potentially model the uncertainty in viewers' initial belief ($\sigma 1$) as a function of the variances in those ratings.

Before a movie is available in the market, a consumer's prior belief of the unobservable appeal is affected by both $B_j$, the intrinsic perception of the unobservable appeal of the movie $j$, and $\Delta A_j^u$, the belief of the unobservable appeal shaped by other factors such as advertising and 1st stage performance. It is well known that advertising can have both informative and persuasive effects. Therefore, it is likely to influence consumers' belief about a movie's unobservable appeal before watching it. As to the platform movie, since it has been in limited markets, by the time of its national release, a consumer may get information indicating its unobservable appeal from many sources which we can reasonably assume to be correlated with its 1st stage box office performance. For example, consumers are likely to hear some good thing about a platform movie prior to its national release if its 1st stage box office is higher. Such positive information about the movie could be in the form of opinions from a consumer's social network, online blogs, media mentioning, etc. Hence, the 1st stage performance of a platform movie can affect consumers' belief of the unobservable movie appeal. Consequently, the 1st stage performance of a platform movie, as well as advertising, can influence the total box office of a movie after its national release through the Bayesian learning process, which, as we discussed earlier, is determined by the word of mouth information received after national release and consumers' prior belief on unobservable movie appeal.

We model $B_j$ as coming from two sources, consumers' past consumption experience with or knowledge about similar movies and, second, information about the unobservable appeal for the specific movie considered. We model $B_j$ as a weighted average between two such information sources, that is,

$$B_j = a_j \overline{A_j^u} + (1 - a_j) A_j^u$$

where $\overline{A_j^u}$ is average unobservable appeal of the movies of the same genre in the sample, which captures consumers' prior knowledge of similar movies,[9] that is, $\overline{A_j^u} = \frac{1}{H} \sum_{k \in G_j} A_k^u$, where $G_j$ is the set of movies with the same genre as movie $j$ and $H$ is the measure of the number of movies in $G_j$. Such an approach is similar to Byzalov and Shachar (2004). Also, $a_j$ is the weight consumers put on their prior experience with movies of genre $G_j$. We constrain $a_j$ to be between 0 and 1 and consider it to vary across genres and release types.

$$a_j = \frac{\exp(\kappa_1 Platform_j + \kappa_2 Genre_j)}{1 + \exp(\kappa_1 Platform_j + \kappa_2 Genre_j)}$$

For $\Delta A_j^u$, we specify it as

[9] Genre is the most straightforward and available variable to measure similarity of movies. We also tried other variables such as MPAA rating and found similar results.

where $Ad_j$ is the cumulative advertising expenditure for movie $j$ before its national release. The interaction between genre and advertising reflects the possibility that the advertising effects may vary across different release types and genres. $BO_{1st,j}$ is the 1st stage box office revenue for a platform movie. For wide release movie, $BO_{1st,j}$ is zero.

### 3.4 Identification

The identification of our model is similar to Ching (2000, 2008) and Narayanan et al. (2005). The identification of the linear parameters in the utility function given the set of nonlinear parameters is straightforward. We first discuss the learning parameters.

The initial prior variance $\sigma_1$ is difficult to identify with our data. With disaggregate data, such a parameter is identified through volatility in purchase decision initially at individual level (Erdem and Keane 1996). However, at the aggregate level, such an effect is smoothed out across individuals which makes it impossible to identify $\sigma_1$. Thus, we fix $\sigma_1 = 1$. This implies that .

The parameters in unobservable appeal $A_j^u$ are identified by market share comparison across movies after the opening week. Suppose there are two movies which have similar prior belief but different unobservable appeals such that one is greater than the prior belief but the other is below it. Both movies will have the same opening week performance. However, the high unobservable appeal movie's market share will decrease more slowly due to positive word of mouth which lifts consumers' perception of its unobservable appeal, while the low unobservable appeal movie's market share will decrease faster due to negative word of mouth. The difference between the market shares across these two movies identifies parameters in $A_j^u$ after controlling for learning rate.

The parameters in prior belief $a_j\overline{A_j^u} + (1 - a_j)A_j^u + \Delta A_j^u$ are identified through the correlation between variation in the opening week box office across movies and variation in terms of indicators for average unobservable appeal $\overline{A_j^u}$, unobservable appeal $A_j^u$ and deviation in unobservable appeal $\Delta A_j^u$. Finally, the uncertainty in word of mouth signals $\sigma_\rho$ is identified from the learning rate, in other words, the inter-correlation between market shares over time.

The unique feature of our data set that relates to the identification of model parameters is that it contains a large number of movies produced every year but each of them has a short life cycle. Therefore, additional information has to be obtained from cross-sectional differences across movies for model identification.

One potential concern of our model is that the release strategies are endogenously determined, that is, movies will be released differently according to certain characteristics. A fully structural model with both release decision and demand function modelled will certainly solve this problem, but it is technically difficult given the information available to us. In our model, this potential endogeneity problem is alleviated as we allow the two release types to lead to different prior choice.[10]

### 3.5 Likelihood and estimation

Let $\delta_{ijt}$ be the mean utility of the purchase option for consumer $i$, movie $j$ at time $t$,

$$\delta_{ijt} = E\left(A^{u}_{ijt}\right) + A^{0}_{ijt}$$

For consumer $i$, the probability she is going to see movie $j$ at time $t$ conditional on her having not seen the movie is

$$\Pr(b_{ijt} = 1, b_{ijr} = 0, \quad \forall r < t) = \left(\frac{e^{\delta_{ijt}}}{1 + e^{\delta_{ijt}}}\right) \prod_{r<t} \left(\frac{1}{1 + e^{\delta_{ijt}}}\right)$$

where $b_{ijt}$ is an indicator of purchase decision. Integrating over the population, we get the expected choice probability for movie $j$ at time $t$,

$$\pi_{jt} = \int \Pr(b_{ijt} = 1, b_{ijr} = 0, \quad \forall r < t) \mathrm{d}f(\xi_{ij1}, \eta_{jr}, \beta_{i10}, \beta_{i11})$$

For each movie, we observe weekly box office revenue from the release week until the movie is dropped from theatres. As stated earlier, since the price of a movie ticket is relatively constant over time, we can project $n_{jt}$, the weekly number of consumers watching movie, from the box office revenue. The observed data $n_{jt}$ are the multinomial outcomes from a process whose probabilities are given by $\pi_{jt}$. Suppose for movie $j$ we have data from time 1 to time $T$. As in Pakes (1986), the likelihood function based on observed aggregate data is given by

$$L(\Omega) = \prod_{j=1}^{J} \left\{ \prod_{t=1}^{T} \pi_{jt}(\Omega)^{n_{jt}} \right\} \left\{ 1 - \sum_{t=1}^{T} \pi_{jt}(\Omega) \right\}^{M - \sum_{t=1}^{T} \pi_{jt}(\Omega)}$$

where M is the measure of market size. Here we define potential consumers for a movie as those who go to see a movie at least once a year. According to MPAA, the total number of moviegoers in 2003 is 167.7 million. We use this number as our market measure for wide release. For platform movies, we use this number less the first stage demand as the market measure for the national release stage.

The model is estimated using simulated maximum likelihood method as discussed in Pakes (1986) with 1,000 random draws. For estimation purposes, we divide the market size $M$, weekly moviegoer $n_{jt}$, and cumulated moviegoer $N_{jt-1}$ by 100 million. Notice that this is purely a scaling effect because equation (6) can be written as

[10] Heckman's two-step approach (with the release decision modelled in a reduced form then controlling for selection bias in the demand function) is difficult to implement here because this requires additional exogenous information uncorrelated with movie demand, which is hard to obtain and justify.

$$\omega_{ijt} = \qquad 1 + \frac{N_{jt-1}}{\sigma_\rho^2}$$

We can see that $n_{jt}$ and $N_{jt-1}$ always go with $\sigma_\rho^2$ in the form of $\frac{n_{jr}}{\sigma_\rho^2}$. Hence $\sigma_\rho$ is scaled down by $10^4$.

## 4 Results

### 4.1 Goodness of fit and model comparison

To get a sense of the fit of our model, we calculate the model prediction for weekly box office revenue (in millions of dollars) for each movie. The difference between weekly box office predictions and actual observations is squared and averaged, which yields $MSE = 11.87$. As the variance in the observed weekly box office revenue is 71.87, this indicates that our model explains about 83 % of the variance in the data.

We also compare our proposed model with a simplified model in which consumers generate initial prior belief (Eq. 6), but there is no learning from WOM signals. This is the same as assuming $\sigma_\rho \to +\infty$. Then given $\gamma_{j1} = \sigma_\rho^2$, Eq. 6 becomes

$$\omega_{ijt} = \left(B_j + \Delta A_j^u\right) + \xi_{ij1}$$

Table 2 reports the model comparison for the full estimation sample. Our model does better according to both BIC and AIC, even though it has more parameters. This suggests that it is important to model the learning effect in order to understand the impact of platform release on demand.

We also validate the analysis in estimation versus holdout sample fashion and compare our model with alternative models. We randomly choose 75 % of the movies as an estimation sample and the remaining 25 % as the holdout sample. We first estimate both models using the estimation sample then use the estimated model to predict weekly box office revenue in the holdout sample. We report the model comparison in Table 2. Note that the proposed model again performs better than the competing model in the estimation sample in terms of BIC, AIC and MSE. Furthermore, in the holdout sample, the proposed model has a better prediction power, which is indicated by a lower MSE.[11]

### 4.2 Demand estimates

We present our results for the model estimation in Table 3. We focus first on parameters in unobservable appeal. Both critics and viewer ratings are positively

[11] The results reported hereafter are based on the estimation that uses the full sample as keeping a holdout sample is no longer necessary.

| | | |
|---|---|---|
| Full sample | | |
| No. of parameters | 39 | 38 |
| No. of observations | 7,471 | 7,471 |
| Log likelihood | −2,472.4 | −2,478.8 |
| AIC | 5,022.8 | 5,033.6 |
| BIC | 5,292.6 | 5,296.5 |
| Holdout sample test | | |
| Estimation sample | | |
| No. of parameters | 39 | 38 |
| No. of observations | 5,616 | 5,616 |
| Log likelihood | −1,854 | −1,859.3 |
| AIC | 3,786.0 | 3,794.6 |
| BIC | 4,044.7 | 4,046.7 |
| MSE | 11.91 | 13.13 |
| Holdout sample | | |
| MSE | 12.44 | 13.29 |

related to the movie unobservable appeal, with viewer rating being the stronger indicator. The total number of Oscar nominations is a strong indicator of unobservable appeal, while the Oscar awards have a positive, but insignificant effect. The production budget has a U-shaped relationship with movie unobservable appeal, but its effect is relatively small. This limited effect may occur since we already control for unobservable appeal through such measures as viewer and critics ratings and awards won.

Next we look at the weight consumers put on their prior experience when generating their prior belief. In particular, consumers put 19 % (calculated as $\frac{\exp(-1.4243)}{1+\exp(-1.4243)}$) and 33 % of weight on prior experience for drama and comedy respectively, but the weight is less than 4 % for action movies. Although we expected that a platform release will help consumers better discern the actual unobservable appeal, which implies less weight on the prior experience, we did not obtain a significant estimate on this parameter.

For $\Delta A_j^u$, our results show that first stage performance enhances consumers' expected unobservable appeal for a platform movie by the time of national release. This suggests that the more successful the movie in the first stage, the higher its perceived unobservable appeal. In terms of advertising effects, the difference between platform and wide release is not significant. Although advertising always shifts consumers' expected unobservable appeal of a movie upwards, its effect differs across genres, as expected. Specifically, advertising is most effective for action movies, followed by comedy, and is least effective for drama. This is consistent with the view that in a 30-second ad, action and comedy movies attract people more than drama, which tends to have more complicated stories.

| | Estimate | SE |
|---|---|---|
| Unobservable appeal | | |
| Intercept | −9.0755** | 0.1373 |
| Oscar_N | 0.0668** | 0.0171 |
| Oscar_W | 0.0256 | 0.0409 |
| Rating_Critics/100 | 6.7382** | 0.6892 |
| Rating_Viewer/100 | 19.7418** | 0.5467 |
| (Production budget/100 millions) | −0.4664 | 0.2655 |
| (Production budget/100 millions)$^2$ | 0.3026 | 0.1759 |
| Weight on prior experience | | |
| Platform | −0.0490 | 0.3036 |
| Drama | −1.4243** | 0.5582 |
| Action | −3.2841** | 0.3647 |
| Comedy | −0.7049 | 0.5766 |
| Deviation in perception | | |
| Ad*Genre/(100 millions) | | |
| Ad*Platform | −0.5924 | 0.7142 |
| Ad*Drama | 0.3163 | 0.4141 |
| Ad*Action | 1.0328** | 0.3421 |
| Ad*Comedy | 0.9314** | 0.3457 |
| 1st stage box office/(1 million) | 0.1400** | 0.0382 |
| LN($\sigma_\rho$) | −4.6200** | 0.3455 |
| Observable appeal | | |
| Major distributor | 0.0911 | 0.0467 |
| Drama | −1.0781** | 0.2452 |
| Action | −0.3552 | 0.2339 |
| PG-13 | 0.1359* | 0.0536 |
| R | 0.0064 | 0.0630 |
| Sequel | 0.1155* | 0.0582 |
| Age*Platform | −0.0712** | 0.0250 |
| Age*Drama | −0.0677** | 0.0123 |
| Age*Action | −0.0883** | 0.0164 |
| Age*Comedy | −0.0527** | 0.0105 |
| Theatre/1,000 | 1.1645** | 0.0336 |
| June | 0.1706* | 0.0619 |
| July | 0.2941** | 0.0587 |
| August | 0.2386** | 0.0559 |
| NEW YEAR'S DAY | 0.4573** | 0.1125 |
| MEMORIAL DAY | 0.3378** | 0.1042 |
| INDEPENDENCE DAY | 0.1275 | 0.1145 |
| LABOR DAY | 0.0958 | 0.2295 |
| THANKSGIVING | 0.1083 | 0.2216 |

| | Estimate | SE |
|---|---|---|
| CHRISTMAS | 0.3857 | 0.2003 |
| Heterogeneity_Drama | −1.5153** | 0.1741 |
| Heterogeneity_Action | −1.1569** | 0.2547 |
| Log likelihood | −2,472.4 | |

* significant at 5 %; ** significant at 1 %

It is interesting to discuss the trade-off between advertising and first stage performance. Given that the average 1st stage box office revenue for a platform movie is $2.42 million, our results reveal that it shifts consumers' expected unobservable appeal to the same extent as a $107.11 million advertising expenditure for dramas, $36.38 million for comedies, and $32.80 million for action movies.[12] This may explain why the majority of the platform movies in our sample are dramas. These economic calculations for the trade-off between advertising budget and first stage performance reveal a possible motivation for using a platform release for movies with small advertising budgets. However, there are other effects that need to be considered as well. These include the finding that platform movies decay more rapidly than wide release movies when they reach nationwide release as we will discuss below.

The standard deviation in the word of mouth signal $\sigma_\rho$ is $\exp(-4.6200) = 0.0100$, which is significant. We will show the effect of learning in a later section.

Turning to the linear parameters in observable appeal, using a major distributor is not significantly related to consumers' evaluation (as in Elberse and Eliashberg 2003). Consumers prefer comedies the most, followed by action and drama. Also, there is significant heterogeneity in consumer's preferences for each genre, which reflects consumers' variation in their preferences for certain types of movie. In terms of MPAA rating, PG-13 is the most popular category. Sequel movies are more preferred than other movies, which is not surprising. In terms of the ageing effect of a movie, which captures how fast a movie's attractiveness declines over time, we expected that platform movies would age faster than wide release movies because consumers may have heard about these movies well before they were widely available. Our estimate indeed reveals such a pattern. We also find the ageing patterns differ across genres with action movies declining faster than the other genres. The theatre coefficient is significantly positive, which confirms that availability increases consumers' likelihood of viewing a movie. Finally, the summer months and the New Year and Memorial Day holidays have significantly more demand.



[12] That is, $1 million in first stage box office revenues shifts consumers' prior belief of unobservable appeal to the same extent as $44.26 million advertising expenditure for dramas, $15.03 million for comedies, and $13.56 million for action movies.

#### 4.3.1 *Movie unobservable appeal versus prior belief*

It is interesting to see how consumers' prior belief deviates from the actual unobservable appeal. With the demand estimates, we calculate the actual unobservable appeal and prior belief for each movie prior to its national release and depict the two variables in Fig. 4. We find that consumers' prior perception of most movies, nearly 87 %, is higher than the actual unobservable appeal, reflecting the impact of the firm's advertising, and for the platform movies in our sample, the effect of the first stage performance.

#### 4.3.2 *Word of mouth effect*

Our model captures the learning process through word of mouth during movie diffusion. An interesting question to ask is how significant the word of mouth effect is. In our model, the word of mouth signals affect movie demand only after the opening week of national release. It affects movie demand in different directions depending on the gap between consumers' prior belief and the actual unobservable appeal. For a movie for which consumers' perceived appeal is higher than the actual unobservable appeal, word of mouth signals decrease consumers' tendency to see the movie after the opening week. On the contrary, if a high-quality movie is perceived as one with low unobservable appeal, the word of mouth signals actually lift the demand after the opening week. We show such interaction for two movies in Fig. 5a and b. "2 Fast 2 Furious" is a movie in which consumers' prior perception (−7.76) is higher than its actual unobservable appeal (−7.92), while "A Beautiful Mind" has lower prior perception (−5.83) than actual unobservable appeal (−5.49). For both movies, we calculate model predicted weekly box office revenue with or without the learning effect. It shows that "2 Fast 2 Furious" suffers from word of

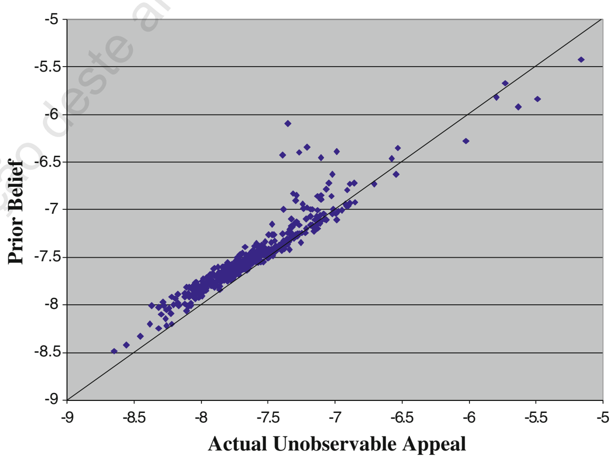


**Fig. 4** Actual unobservable appeal versus prior belief before national release

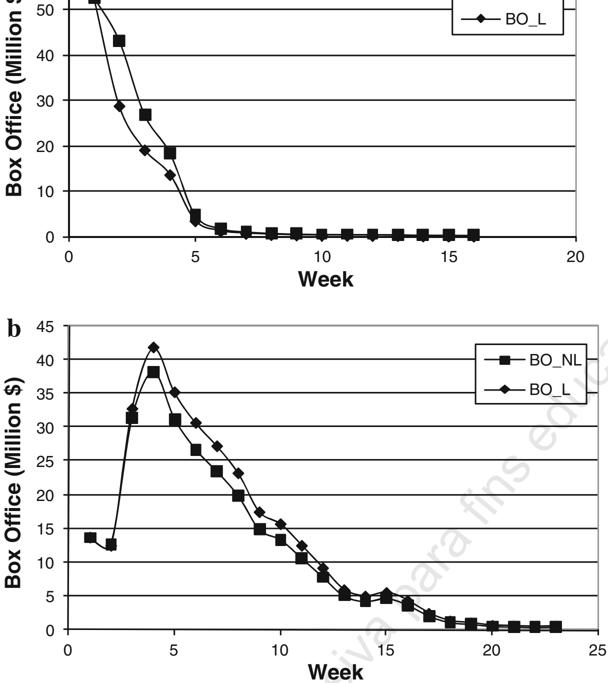


**Fig. 5** **a** Learning effect: 2 Fast 2 Furious. *BO_L* Model predicted weekly box office revenue with learning effect. *BO_NL* Model predicted weekly box office revenue without learning effect. **b** Learning effect: A Beautiful Mind. *BO_L* Model predicted weekly box office revenue with learning effect. *BO_NL* Model predicted weekly box office revenue without learning effect

mouth effect from the second week, while "A Beautiful Mind" benefits from such a learning effect.

We calculate the relative impact of word of mouth on cumulative box office revenue for each movie, $\frac{cumbo_L - cumbo_{NL}}{cumbo_{NL}}$, where *cumbo* is the cumulative box office, $L$ indicates with learning which is as our proposed model, and NL indicates no learning which is denoted as model 1 in the model comparison section discussed above. For each model, we first calculate the weekly box office revenue for each movie given our demand estimates. We then sum up over weeks to get the cumulated box office revenue. We find the impact ranges from −0.60 to 0.12, with mean of the absolute impact $\left|\frac{cumbo_L - cumbo_{NL}}{cumbo_{NL}}\right|$ as 0.26, that is, on average the learning effect makes a 27 % difference in cumulated box office revenue.

## 4.4 Counterfactual analysis: switching platform movies to wide release

An interesting question to ask is what would have happened if studios had used a different release type for the movies in our sample. This can be answered by movies, we focus on examining the effect of switching a platform release to a wide release. We also note that we are not attempting to fully explain why the studios choose a particular release type for a movie. To answer such a question, we would have to develop a complex model to endogenize a studio's release choice, which is beyond the scope of this paper. However, our research does shed some light on the trade-off between the two release types for the box office after national release from both the demand estimation and counterfactual analysis.

In the counterfactual analysis, we compare the profitability of wide release versus platform release for the platform movies in our sample. The details of the decision process and timing of the movie's release in the counterfactual analysis are provided in the Appendix. Based on this process, we calculate the advertising expenditure, distribution and movie life period (length of theatrical run) using the information from wide release movies in our sample. In Table 4, we report both the predicted and the actual observations. The two release types differ significantly in advertising expenditure and distribution coverage, with the wide release accompanied by a higher level in both dimensions.

**Table 4** Counterfactual analysis: change platform movies to wide release

| Scenario | Variable | Mean | SD | Minimum | Maximum | Median |
|---|---|---|---|---|---|---|
| Successful platform movies ($N = 49$) | | | | | | |
| Platform release | Advertising expenditure | 4.57 | 4.19 | 0.02 | 16.2 | 3.31 |
| | Maximum no. of theatres | 997.2 | 683.26 | 224 | 2,701 | 842 |
| | Life | 18.51 | 8.49 | 7 | 48 | 17 |
| | Cumulated box office | 32.61 | 45.43 | 1.59 | 238.59 | 15.51 |
| | Profit (equally split) | 6.86 | 15.74 | −8.33 | 103.27 | 4.60 |
| | Profit (sliding contract) | 6.01 | 16.57 | −8.86 | 77.69 | 1.10 |
| Wide release | Advertising expenditure | 8.13 | 2.71 | 2.57 | 13.78 | 8.19 |
| | Maximum no. of theatres | 2,065.51 | 653.71 | 1,036.00 | 2,877.00 | 1,852.00 |
| | Life | 23.61 | 4.60 | 13.00 | 34.00 | 24.00 |
| | Cumulated box office | 104.41 | 138.01 | 14.73 | 688.65 | 37.95 |
| | Profit (equally split) | 37.88 | 67.12 | −5.54 | 327.67 | 8.40 |
| | Profit (sliding contract) | 27.77 | 50.64 | −5.88 | 245.83 | 5.72 |
| Comparison | ΔProfit (equally split) | 31.02 | 56.76 | −11.75 | 224.40 | 3.17 |
| | No. of success (equally split) | 34 | | | | |
| | ΔProfit (sliding contract) | 21.75 | 46.21 | −75.33 | 193.96 | 2.21 |
| | No. of success (sliding contract) | 36 | | | | |

Advertising expenditure, box office, and profit are all measured in millions of dollars

ΔProfit: Profit under wide release-Profit under platform release

No. of success: Number of movies better off after switched to wide release

We begin by calculating the studios' predicted profits from the platform movies given that they eventually go for national release. We first make certain assumptions on the cost and revenue sharing in the distribution channel. Note that studios incur certain fixed cost for each theatre to play their movies, such as the cost of making a copy of the film. Consistent with the industry data reported in Einav (2007), we assume it costs \$3000 for each theatre. Also, studios often have revenue sharing contracts with theatres. From our review of the relevant literature, we consider two contracts in our analysis: equally split and sliding scale contracts. Under the equally split contract, studios and theatres split the total box office revenue approximately equally. Filson et al. (2005) analyse 2,769 movie exhibition contracts signed by the Wehrenberg movie theatre chain in St. Louis and find that the studios retained 54 % of the gross admissions revenue on average. Besides the equally split contract, we also consider a sliding scale contract in which movie studio's share of revenue decreases over time. Swami et al. (2003) examine a set of actual contract terms faced by an exhibitor in a season and find 60/50/40/35 to be the terms for the modal contract. Here we assume a similar sliding scale contract with terms 60/50/40/35 and 35 after week 4. Since platform movies are released in two stages, we assume the sliding scale contract restarts at the week of national release. Discussions with movie industry executives confirm that these are reasonable assumptions.

Given these assumptions, we first calculate the predicted box office revenue after national release for each platform movie using our demand estimates. Then, we calculate the predicted profit for movies under platform release

$$\prod_{j,platform} = \text{Revenue}(Contrast_j, BO_{j1}, BO_{j2}) - Ad_j - 3{,}000 * Theatre_j$$

where $\Pi_{j,platform}$ is the profit for movie $j$ under platform release, $Contract_j$ refers to the contract employed between the movie studio and the theatres, $BO_{j1}$ and $BO_{j2}$ are the total box office revenue in 1st stage and national release respectively, $Ad_j$ is the advertising expenditure, and $Theatre_j$ is the maximum number of theatres engaged during the period of national release. A movie studio's revenue is a function of its contract with the theatres and total box office revenue. Note that production cost is not in the formula because we consider it as a sunk cost at the time the movie distribution decision is made. We report descriptive statistics for movie box office and profit under platform release in Table 4. Most movies are profitable under our definition with an average studio profit of \$6.86 million under the equally split contract and \$6.01 million under the sliding scale contract.

### *4.4.2 Counterfactual: wide release for successful platform movies*

Next, we conduct a counterfactual analysis to explore the potential outcome of wide releasing the platform release movies in our sample. Given the predicted advertising expenditure and weekly number of theatres, we first calculate the weekly box office revenue under a hypothetical wide release strategy for each platform movie using

$$\prod_{j,wide} = \text{Revenue}(Contrast_j, BO_j) - Ad_j - 3,000 * Theatre_j$$

We report the results of the counterfactual analysis in Table 4. We find that profits increase for 34(36) of the 49 platform movies under an equally split (sliding scale) contract in our sample. This suggests that about 70 % of the movies in our sample could have done better if they had been given a wide release. This indicates there is still an opportunity for the studios to improve their releasing format decisions. The movies with improved profits if they were to be wide released in our counterfactual analysis tend to be the ones with high appeal (their average unobservable appeal is −6.98 compared to −7.28 for the other platform movies). Furthermore, these movies averaged 4.0 nominations for Oscars and won 0.9 of them, more than 10 times the rate for the rest of the platform movies.

## 5 Conclusion

In this paper, we propose a formal learning model to capture consumers' decision to watch a movie. In particular, we detail the process through which the movie release type affects consumers' movie evaluation and subsequent purchase decisions after a movie's national release. Our model is based on both the movie literature and the consumer learning literature, as well as capturing conventional wisdom in the movie industry. As each movie has only a short life cycle, our estimation method incorporates both cross-sectional and times series data. In our empirical analysis on a sample of movies that either had a wide release or a platform release but eventually were given a national release, we find that platform release helps to shift consumer's prior perception of the unobservable appeal of a movie upwards, which is much more effective than advertising. However, platform release tends to accelerate the decay of movie attractiveness. The contrast between the above two effects makes the release type a movie-specific choice for studios planning for an eventually national release of the movie. Using counterfactual analysis, we find that a considerable proportion of platform movies could have done better with a wide release. This suggests that there is room for the studios to improve their movie releasing format decisions.

It is important to point out some limitations of our study, which could be avenues for future research. We highlight two issues here. First, word of mouth can be generated and spread in different ways, and these effects could vary significantly. However, in our model, we simplify them as one format due to data limitations. It would be interesting to investigate the relative impact of different types of word of mouth effects, which could provide useful managerial implications. Second, we did not directly model a studio's choice of release type; this would be a challenging task which requires information on the expectation of the movie performance, which were not available here. This is certainly a research direction worth pursuing.

of Canada and the Greater China Business Research Institute, Cheung Kong Graduate School of Business.

## Appendix: switching platform movies to wide release

In the counterfactual analysis, we compare the profitability of wide release versus platform release for the platform movies in our sample. The decision process of the movie studio is assumed to be as follows. In stage 1, the national release date is set. It is exogenous in our counterfactual analysis, and it is set as the takeoff week of the platform movie. In stage 2, the movie studio decides to adopt a wide release or platform release strategy, given the release date ($T_{0j}$), movie characteristics, $X_j$, and the unobservable (to the studio) appeal of the movie, $\widetilde{A}_j^u$. In stage 3, the movie studio decides on the advertising expenditure ($Ad_j$) and number of theatres engaged in the opening week. Once the movie is released, the number of theatres in the following weeks is largely determined by box office performance (Krider et al. 2005). In addition, we make the following assumptions:

A1 On average, $\widetilde{A}_j^u$ equals the realized actual unobservable appeal of a movie, $A_j^u$. Specifically, we write $\widetilde{A}_j^u \sim N\left(A_j^u, \sigma_{\tilde{A}}^2\right)$.

A2 Observed advertising expenditure, number of theatres, and movie life period are assumed to be the equilibrium outcome given $X_j$ and $\widetilde{A}_j^u$. However, they cannot be determined from the estimated demand model directly because we have no information on such missing factors as the financing ability and risk preference of the studio and the bargaining power of studio versus exhibitors. Instead, we use reduced form regressions to establish the relationships between those equilibrium outcomes and $\left(X_j, \widetilde{A}_j^u\right)$, that is, $\left(X_j, \widetilde{A}_j^u\right)$ are the independent variables in those regressions.

Given the above assumptions, we can compare the profitability of wide release versus platform release for movies given their characteristics ($X_j$, $\widetilde{A}_j^u$, $T_{0j}$. We can assume $\widetilde{A}_j^u = A_j^u$ in our analysis to see whether the observed platform releases are optimal.

### Advertising, Number of Theatres, and Movie Life under Wide Release

Before conducting the counterfactual analysis, we have to predict the advertising budget, the length of the theatrical run of the movie ("life" period), and the number of theatres engaged in each week for a platform movie should it switch to wide release. As discussed above, we use a reduced form approach to establish the link between these variables and $\left(X_j, \widetilde{A}_j^u\right)$, which are movie characteristics and perceived unobservable appeal by the studio.

We assume for a wide release movie, the advertising expenditure is decided by

$$Ad_j = \alpha + \beta_1 X_j + \beta_2 \tilde{A}_j^u + \varepsilon_j$$

where $Ad_j$ is the advertising expenditure for wide release movie $j$. Note that we have no observation on studio's perceived unobservable appeal $\tilde{A}_j^u$. Instead, we know $A_j^u$ through our demand estimation. Therefore, we can write the equation as

$$Ad_j = \alpha + \beta_1 X_j + \beta_2 A_j^u + \chi_j$$

where $\chi_j = (\beta_2 \varepsilon_{\tilde{A}j} + \varepsilon_j)$, where $\varepsilon_{\tilde{A}j} \sim N\left(0, \sigma_{\tilde{A}}^2\right)$. Therefore, $\chi_j$ is still *i.i.d.* normal distribution with mean zero.

*Number of Theatres Engaged*

We define two regressions for the opening week and the following weeks separately. This is because the number of theatres in the opening week is often determined based on expected demand, while in the following weeks, the number of theatres is driven largely by the last week's performance (Elberse and Eliashberg 2003; Krider et al. 2005). Therefore, we have

$$\log(Theatre_{j1}) = \alpha_1 + \beta_{10} X_j + \beta_{11} \tilde{A}_j^u + \beta_{12} Ad_j + \beta_{13} Season_{j1} + \varepsilon_{j1}$$

for the opening week, and

$$\log(Theatre_{jt}) = \alpha_2 + \beta_{21} X_j + \beta_{22} \tilde{A}_j^u + \beta_{23} Ad_j + \beta_{24} \log(BO_{jt-1}) + \beta_{25} Age_{jt} + \varepsilon_{jt}$$

for the following weeks, where $Season_{j1}$ is a vector of month and holiday variables for opening week and $BO_{jt-1}$ is the box office revenue for movie $j$ in the last week. By a similar argument to that for advertising, we rewrite the equations as

$$\log(Theatre_{j1}) = \alpha_1 + \beta_{10} X_j + \beta_{11} A_j^u + \beta_{12} Ad_j + v_{j1}$$

$$\log(Theatre_{jt}) = \alpha_2 + \beta_{21} X_j + \beta_{22} \tilde{A}_j^u + \beta_{23} Ad_j + \beta_{24} \log(BO_{jt-1}) + \beta_{25} Age_{jt} + v_{jt}$$

where $v_{j1} = (\beta_{11} \varepsilon_{\tilde{A}j} + \varepsilon_{j1})$ and $v_{jt} = (\beta_{22} \varepsilon_{\tilde{A}j} + \varepsilon_{jt})$, which are *i.i.d.* normal distribution with mean zero.

*Movie life period*

The length of a movie's theatre run under wide release is defined according to the number of theatres engaged. We assume that a movie is out of the market when the predicted number of theatres is less than 95, which is the mean of the observed number of theatres in the last week across all movies in our initial sample.

We estimate the above regression equations using observations of the wide release movies in the sample and therefore predict those variables for the platform movies. In addition, we restrict the number of theatres engaged in the opening week

## References

Anslie, A., Drèze, X., & Zufryden, F. (2005). Modeling movie life cycles and market share. *Marketing Science, 24*(3), 508–517.

Basuroy, S., Desai, K. K., & Talukdar, D. (2006). An empirical investigation of signaling in the motion picture industry. *Journal of Marketing Research*, *43*(2), 287–295.

Byzalov, Dmitri., & Shachar, Ron. (2004). The risk reduction role of advertising. *Quantitative Marketing and Economics, 2*, 283–320.

Ching, A. T. (2000). *Dynamic equilibrium in the U.S. prescription drug market after patent expiration*, Ph.D. Dissertation, University of Minnesota.

Ching, A. T. (2008). *Consumer learning and heterogeneity: Dynamics of demand for prescription drugs after patent expiration*. Working Paper, University of Toronto.

DeGroot, M. H. (1970). *Optimal statistical decisions*. New York: McGraw-Hill.

Einav, L. (2007). Seasonality in the U.S. motion picture industry. *Rand Journal of Economics, 38*(1), 128–146.

Elberse, A., & Eliashberg, J. (2003). Demand and supply dynamics for sequentially released products in international markets: The case of motion pictures. *Marketing Science, 22*(3), 329–354.

Eliashberg, J., & Sawhney, M. S. (1994). Modeling goes to hollywood: Predicting individual differences in movie enjoyment. *Management Science, 40*, 1151–1173.

Erdem, T., & Keane, M. P. (1996). Decision-making under uncertainty: Capturing dynamic brand choice process in turbulent consumer goods market. *Marketing Science, 15*(1), 1–20.

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

Kadiyali, V., Sudhir, K., & Rao, V. R. (2001). Structural analysis of competitive behavior: New empirical industrial organization methods in marketing. *International Journal of Research in Marketing, 18*(1/2), 161–186.

Krider, R. E., Li, T., Liu, Y., & Weinberg, C. B. (2005). The lead-lag puzzle of demand and distribution: A graphical method applied to movies. *Marketing Science, 24*(4), 635–645.

Krider, R. E., & Weinberg, C. B. (1998). Competitive dynamics and the introduction of new products: The motion picture timing game. *Journal of Marketing Research, 35*(1), 1–15.

Narayanan, S., Manchanda, P., & Chintagunta, P. (2005). Temporal differences in the role of marketing communication in new product categories. *Journal of Marketing Research, 42*(3), 278–290.

Neelamegham, R., & Chintagunta, P. (1999). A Bayesian model to forecast new product performance in domestic and international markets. *Marketing Science, 18*(2), 115–136.

Nelson, P. (1974). Advertising as information. *Journal of Political Economy, 82*, 729–753.

Pakes, A. (1986). Patents as options: Some estimates of the value of holding European patent stocks. *Econometrica, 54*, 755–784.

Sawhney, M. S., & Eliashberg, J. (1996). A parsimonious model for forecasting gross box-office revenues of motion pictures. *Marketing Science, 15*(2), 113–131.

Swami, S., Lee, E., & Weinberg, C. B. (2003). Channel contracts in the movie industry: Show me the money. *Proceedings from the Third and Fourth Business and Economics Scholars Workshop in Motion Picture Industry Studies,* Florida Atlantic University, pp. 65–76.



**Fonte: Journal of Cultural Economics, Holanda. Vol. 37, n. 3, p. 359-386, 2013.**