Social Learning and Optimal Advertising in the Motion Picture Industry

Hailey Hayeon Joo*
Department of Economics
Ohio University

First Version: November 27, 2007
This Version: February 15, 2009

Abstract

Social learning is thought to be a key determinant of the demand for movies. This can be a double-edged sword for motion picture distributors, because when a movie is good, social learning can enhance the effectiveness of movie advertising, but when a movie is bad, it can mitigate this effectiveness. This paper develops an equilibrium model of consumers' movie-going choices and movie distributors' advertising decisions. First, we develop a structural model for studios' optimal advertising strategies, taking into account the expected social learning process, and a model for consumers' movie demand, given an initial indicator of movie quality (critic ratings) as well as an initial level of advertising. Consumers are assumed to be initially uncertain about movie quality. This, however, is resolved over time through Bayesian updating. That process depends upon (1) the number of previous viewers and (2) their ratings reported over the Internet. We then estimate the model parameters using data pertaining to 236 movies that were shown in theaters in the U.S. between January 1, 2002 and December 31, 2003. The empirical results show that social learning has a positive multiplier effect on movie advertising, with the multiplier effect being strongest for good movies. The simulation of the effects of social learning relative to a world without such learning shows that for good movies, producers spend substantially more on advertising when there is learning involved than they would if there were no learning. For bad movies, social learning makes much less difference to the level of advertising expenditures. Thus, the studio's advertising spending is sensitive to both consumer uncertainty about movie quality and the speed with which potential movie-goers learn about movie quality.

Keywords: Social Learning, Advertising, Motion Pictures
JEL Classifications: D83, L82, M37

---

*I am truly indebted to my advisors Petra Todd, Aureo de Paula and Joel Waldfogel for their invaluable advice and encouragement. I have greatly benefited from discussions with Roy Boyd, Flavio Cunha, Elena Krasnokutskaya, Donghoon Lee, Antonio Merlo, Janet Pack, Kevin Song and Kenneth Wolpin. I thank the NSF for financial support, through the Doctoral Dissertation Improvement Grant in Economics, #0649483. I am also grateful to Amy Lamare at the Hollywood Stock Exchange, who made possible my access to the data. All errors are mine. Email: joo@ohio.edu.

1 Introduction

Social learning is thought to influence people's movie-watching behavior, the decisions about whether and when to watch a film.[1] Like other experience goods, the quality of a movie is difficult to assess in advance and is determined only upon consumption. Through social learning, potential movie-goers learn about the quality of a movie from people who have watched it. For example, critics' ratings, word-of-mouth referrals and box office performance may offer potential movie consumers some exposure to the quality of a movie prior to making their choices. These social learning mechanisms play a particularly important role in the motion picture industry.

Although weekly box-office revenue exhibits a general decay pattern, attendance patterns vary substantially by movie. Movies can be categorized into two broad groups. Figure 1 shows the decay patterns of two films, which are representative of these groups. The first group consists mostly of blockbuster movies, such as “Catch Me If You Can” (2002). These films exhibit a convex decay pattern: because their weekly box-office performances begin at a high level, but attendance decreases sharply, eventually smoothing out and slowly fading. The second group consists of the so-called flops, such as the action thriller “Ballistic: Ecks vs. Sever” (2002). Their box-office performances begin at a relatively low level, and the early audience decay is not as steep as that of popular movies. In general, these movies also have a shorter life-span. This pattern reflects not only a failure to capture early interest but also an accelerated audience decline, possibly due to a bad reputation. In this paper, we suggest that these two patterns are intrinsically related to social learning and that both of them can be explained using a similar mechanism.[2]

Social learning then can be thought of as a double-edged sword for motion picture distributors. Once a film is produced, the predominant way to boost demand for it is through advertising. How-

[1]Table 1 shows decision factors to see a movie. 19% of survey respondents answer that they consider others' recommendations, and 18% report that they consider good reviews.

[2]De Vany and Walls (1996) suggest that there is positive information feedback among film audiences that is captured by a dynamic updating process. They show that, without social learning, the relationship between the log of a film's weekly box-office revenue and the log of its rank among other movies playing that week should follow a linear relationship.

ever, the social learning channel can amplify or impede the effectiveness of advertising, depending on the movies' quality and the degree of uncertainty about their quality. Social learning can enhance the effectiveness of advertising for movies that have been positively reviewed. However, if a movie garners bad reviews, learning can mitigate the effectiveness of advertising.

This paper develops an equilibrium model describing consumer decisions about whether and when to watch a movie as well as movie distributor decisions about how much to spend on advertising at the time of the movie's release. The demand model assumes that consumers are myopic, repeatedly making decisions over time about whether to see a movie if they have not seen it before. Studios forecast the stream of movie demand in the future by anticipating social learning effects and use their forecasts to make a decision about their total advertising spending at the time of a movie's release. The structural model developed here describes studios' optimal advertising strategies, taking into account the expected social learning process, and consumers' movie demand, given an initial indicator of movie quality (critic ratings) and an initial level of advertising. Consumers are assumed to have uncertainty about movie quality. Such uncertainty, in turn, is resolved over time through Bayesian updating. This process depends on the number of previous viewers and their ratings reported over the Internet. We estimate model parameters using data pertaining to 236 widely released movies that were shown in theaters in the U.S. between January 1, 2002 and December 31, 2003.[3] This estimation is based on data from multiple sources, including aggregate data on box office performance, census data on household characteristics, survey data on movie consumption, information on motion picture theater prices, critics' ratings of movie quality, Internet ratings of movie quality, information about movie characteristics, and *Hollywood Stock Exchange* (HSX) data on expected box office performance.

The model is estimated by the Simulated Method of Moments (SMM). To evaluate the out-of-sample fit by the model, we randomly leave out 36 movies at the estimation stage and then assess how well the estimated model forecasts the decay patterns for those movies.[4] The within-sample fit and out-of-sample fit of the model are found to be reasonably good, although the fit is better

[3]Here, we define a widely released movie as one having been screened in at least 600 theaters.

[4]To select 36 random movies for exclusion, we ascribe numbers from 1 to 236 to movies in our dataset, generate 36 random numbers uniformly using MATLAB and exclude the movies whose numbers are picked by the program.

for lower-rated movies with strong estimated social learning components. The model's forecasts for the first four weeks of box office revenues turn out to be quite similar to those implied by films' HSX prices on their release dates, which were not used in estimating the model.

After estimating the model, we then use it to explore how changes in advertising influence box office performance, dynamic social learning and movie-going patterns, and ultimately the revenue stream of each movie. The results indicate that, for the most part, social learning has a positive multiplier effect on movie advertising spending. Not surprisingly, for good movies, uncertainty about movie quality leads people to postpone their decision to see a movie relative to the situation in which everyone immediately knows the quality of a movie. Advertising campaigns can induce people to see the movie early on. As more people see it, word-of-mouth referrals about the film's good quality disseminate quickly. Hence, the multiplier effect of social learning on advertising can be strong for good movies. For bad movies, advertising can, to some extent, take advantage of people's ignorance about quality and induce them to see a movie early on, when the uncertainty is greatest. However, its effectiveness is mitigated by the fact that advertising leads to word-of-mouth referrals and quickly disseminates the information that the movie is of poor quality. Thus, for bad movies, the two effects of advertising go in opposite directions, and it remains an empirical question as to which one of these dominates. Our empirical results show that movie distributors' advertising expenditures are highly sensitive to the quality of the movie and that movie studios advertise more heavily for good movies. The simulation of the effects of social learning relative to a world without learning shows that for good movies, producers spend substantially more on advertising in a world with learning than they would in one without it. For bad movies, social learning makes less difference to the level of advertising expenditure. Simulations from the model show that if there were less uncertainty about movie quality, bad movies would be less heavily advertised.

The rest of the paper is organized as follows: Section 2 reviews the related literature. Section 3 describes the data and the industry background. Section 4 discusses the relationship between social learning and advertising. Section 5 describes the model and its solution. Section 6 presents the estimation method and Section 7 shows the results. Section 8 suggests the policy implications, and Section 9 concludes.

2 Review of the Literature

The motion picture industry has recently been the focus of various empirical industrial organization studies. The consequences of movie consumers' social interactions and the effects of movie studios' advertising decisions in particular have been a subject of great interest from economists. Most empirical work based on this literature has not dealt with both the demand-side and the supply-side factors together, but rather has focused on only one of them alone. This paper, however, analyzes the demand for movies and the supply of movie advertising within a single framework. By so doing, we build upon the previous literature of social learning process and the growing body of work which examines advertising's effects in the movie industry.

A series of papers by De Vany and Walls (1996), Moul (2007), Santugini (2006) and Moretti (2008) focuses on the effects of social learning upon aggregate movie demand. The results of their studies confirm the existence of social learning and suggest ways to combine such social learning mechanism into the movie demand model.

Social learning has been incorporated into economic theory during the last 10 to 20 years, but its importance has only rarely been demonstrated empirically. The movie industry, however, offers unique setting to observe this phenomenon. De Vany and Walls (1996) demonstrate the presence of the existence of a social learning mechanism in both their theoretical and empirical models. First, in terms of theory, they explain social learning using a dynamic information-updating process for a movie's quality. They then provide that the empirical rank-revenue relation does not have a Pareto distribution, using their sample of 300 movies collected from *Variety*'s Top-50 between 1985 and 1986. They suggest that this relationship implies a social learning mechanism. They next show that the relationship between the log of a film's weekly box office revenue and the log of its rank among other movies playing that week follows a concave relationship. According to Ijiri and Simon (1974), the deviations in the concave direction from a log-linear rank-revenue relationship imply autocorrelated growth in movie attendances.[5] This autocorrelated growth in movie consumption can be explained by word-of-mouth effects, since people usually go to a movie only once. De

[5] Ijiri and Simon (1974) show that firms have autocorrelated growth when the log-linearized ranks and the log-linearized revenues has a concave downward relationship.

Vany and Walls (1996) provide the present evidence that weekly revenues are autocorrelated and illustrate how social learning mechanisms can exist in the theatrical market.

Recently, Moul (2007) has proposed a method for identifying social learning by looking at whether weekly box office revenue exhibits autocorrelated growth, (*i.e.,* residuals in weekly movie demand are seasonally correlated) and shows that autocorrelated growth in movie consumption can be explained by social learning's effects, since a movie is less likely to be repeatedly consumed. He offers a model of demand that can accommodate both saturation and social learning, and he then considers its implications within an error components framework. Applying this technique to U.S. aggregate box office sales of 1252 movies obtained from *Variety*'s Top-50 between 1990 and 1996, his estimates suggest that (1) social learning is statistically and economically significant and that (2) information about movie quality transmits quickly to the average movie consumer in commonly observed situations. He then provides simulations using these estimates which seem to confirm that social learning can influence on how movies succeed or fail in the theaters.

Another recent paper by Santugini (2006) proposes a model for movie demand and provides empirical estimates which support the existence of social learning within the movie industry. He derives the weekly market share for a movie using a dynamic discrete choice model in which consumers may watch a particular movie at most once and learn about its quality from observing its market share during the release week, ignoring critics and viewers' ratings. Next, using US market-level data of 1007 movies released (between 1998 and 2006 available at variety.com and boxofficemojo.com), he estimates the structural parameters of demand via the maximum simulated likelihood (MSL) procedure. He constructs a distribution for private information signals influenced by other viewers and estimates its mean and variance through the model. By showing that these estimates are statistically significant, he argues that social learning can exist in the movie industry.

Most recently, Moretti (2008) proposes that weekly decay pattern of movie demand is consistent with social learning process. He investigates whether the consumption decisions of individuals depend on information they receive from their peers when product quality is difficult to observe in advance. Focusing on situations where quality is ex-ante uncertain and consumers hold a prior on quality (which they may update based on information from their peers), he argues that this information may arise from the observation of peers' purchasing decisions. With the assumption that every individual receives an independent signal of the goods quality, he proposes that the purchasing decision of one consumer provides valuable information to other consumers, as individuals use the information contained in others' actions to update their own expectations on quality. Using aggregate box-office data of 4,992 movies released between 1982 and 2000 from ACNielsen-EDI, he tests the implications of a simple model of social learning in which the weekly box office revenues are related to the information about their peers' movie consumption as well as movie characteristics including advertising expenditures. He shows that unexpected success or failure in the box office gross during a movie's opening week diverges over time. Furthermore, this divergence is small for movies in which consumers have strong priors and large for movies in which consumers have weak priors. He explains that the effect of a surprise is stronger for audiences with large social networks, and he suggests that social learning appears to be an important determinant of sales in the movie industry. These findings, in turn, imply the existence of a so called "social multiplier" (*i.e.*, the elasticity of aggregate demand to movie quality is greater than that of individual demand to quality).

These authors all show that social learning affects consumers' movie demand by detailing the relationship between certain variables, residuals and past movie demands. They do not, however, consider a way to measure the perceived qualities of movies or how to characterize the social learning process using viewers' ratings. Moreover, none of them consider how social learning can affect movie distributors' optimal levels of advertising.

The effects of advertising on movie demand have been investigated empirically in Moul (2006), Elberse and Anand (2007), and Wilbur and Rennhoff (2007). The focus of these three studies is on the effect of additional advertising per week on movie consumption.

Moul (2006) proposes that the omission of movie advertising can cause substantial biases in estimation of movie demand. He shows that, despite marginally significant impacts on demand, the omission of advertising from movie demand can substantially distort inferences, especially regarding the impacts of other endogenous variables. Using Monte Carlo experiments, he shows that that omitting advertising in demand biases price coefficients downward, (i.e., causing them to be estimated as more elastic than they actually are). To consider the impact of weekly advertising, he measures advertising by the size of newspaper ads for movies released between 1990 and 1996 from the *Sunday Chicago Tribune*.

Elberse and Anand (2007) examine how changes in weekly advertising affect the expectation update about a movie's cumulative box office gross using the transactional data from the HSX market. Although analysis of the returns to advertising is important to the understanding of why advertising spending continuously increases, it is hard to disentangle the causal effect of advertising on sales using data on actual box-office receipts. For this reason, they use the HSX market data for expectation update of a movie's success and the *Competitive Media Reporting* (CMR) data for weekly advertising spending. Their dataset covers 280 movies released between 2001 and 2003. They show that advertising has a positive and statistically significant impact on market-wide expectations prior to release, and that this impact of advertising is lower for movies of lower quality. Their estimate, in turn, implies that the return to advertising for low-quality movies is negative.

Rennhoff and Wilbur (2007) estimate the effect of changes in weekly advertising on box-office revenue. They examine how the weekly post-release advertising spending influences the demand for a movie or changes the expectation for that movie's success. For this estimation, they use the box-office figures from boxofficemojo.com and the advertising spending data from TNS Media Intelligence/CMR for 29 movies released during four weeks of spring, 2003. To avoid the simultaneity problem involved by looking at the weekly advertising decision and the weekly ticket sales, they control for the effects of omitted variables using an instrumental-variables approach. They suggest that post-release movie advertising is profitable, especially when compared with previous findings of pre-release advertising profitability.

They do not consider, however, how the social learning process can influence the effectiveness of advertising. Also, although post-release advertising decisions are important, most of the advertising outlays for a given movie are determined at the beginning of a movie's release. For this reason − and also for reasons of tractability − this paper studies the determinants of an initial advertising budget and how it can influence the overall demand for a film.

3 Data and Background

3.1 Data

The dataset used here covers movies widely released in theaters in the U.S. between January 1, 2002 and December 31, 2003. For each movie, the dataset includes general information,[6] theatrical market information, and movie consumer information.

The data for the distributor, genre, MPAA rating and total running time were obtained from *Variety*. Critic ratings were acquired from Metacritic. The user ratings were collected from the IMDb.[7] The production budgets were obtained from several Internet movie websites, including the IMDb and the websites *Box Office Mojo* (www.boxofficemojo.com), *The Numbers* (www.the-numbers.com) and *Box Office Prophets* (www.boxofficeprophets.com).

The theatrical market information consists of official release dates in the theatrical market, total box office revenue, weekly box office revenues, weekly number of screens, yearly average ticket price and the total advertising expenditure for the theatrical market. These data were also obtained from *Variety*, with the exception of the annual average ticket price, which was obtained from *Box Office Mojo*, and the total advertising expenditure for the theatrical market which came from *Brandweek's Superbrand Report*.[8]

The data for consumers are derived from the *Simmons National Consumer Survey* (NCS) 2002 fall study and the *Integrated Public Use Microdata Series* (IPUMS) consumer survey. They are based on consumption surveys examined by 10,920 households. The NCS fall data set does not provide precise information about the propensity of each household to consume movies. Rather, it includes data on how many households belonging to a given income interval watch specific movies. Also, the movies covered by the NCS 2002 fall set are different from the movies we focus on here. Nevertheless, if the parameters are constant over time, we can estimate consumer attributes from the NCS 2002 fall set. To use the NCS data, we divide the population into both three bread groups

[6]The general information consists of the distributor, genre, MPAA rating, total run time, critic rating, user rating, Academy Award nomination, Academy Award receipt and an estimate of the production budget.

[7]Table 3 shows the summary statistics for Metacritic and IMDb ratings of movies in our data set.

[8]Table 4 shows the summary statistics of important movie characteristics in our data set.

with respect to household income and three groups according to the age of the youngest child in the household. For the three income groups, we permit the marginal utility of income across groups to be different, and for the three groups categorized by the age of the youngest child, we allow the average consumption probability for family-genre movies to vary. For the simulation and estimation of households' movie demands, we use the data from the IPUMS, which is based on a large random sample of U.S. households and provides detailed information about each household. Using the NCS and the IPUMS data together, it is possible then to construct the moment conditions related to household movie consumption.

The initial sample consists of 278 movie titles. However, the dataset does not include the total advertising expenditures for 24 of those titles. Furthermore, we could not observe estimates of the production budgets for 18 movies. Thus, in total, our final data set is comprised of 236 movie titles.

3.2 Background

Social Learning

Social learning, as it applies to economic theory, is a process by which consumers' experiences and information disseminate to other potential consumers. Social learning may include: (1) observing others' choices of a particular good or service; or (2) receiving direct recommendations based on others' first-hand experiences. Hence, the social learning effect can be distinguished from the effect of advertising. The first types of social learning convey viewers' consumption information. Although any single trip to a movie does not by itself imply that the movie is of high quality, the overall popularity as measured by total consumption can indeed suggest high quality. Similarly, the unpopularity of a given movie can imply that it is of lower quality. For this reason, the cumulative market share can be regarded as a variable capturing the so-called *bandwagon effect*. The second type of social learning involves movie consumers' private knowledge of the experience disseminated to potential flim-goers, which is based on the information updating process. Throughout this paper, consumers are assumed to have uncertainty about movie quality. In the opening week of a movie, the review of film critics are the most reliable information available to consumers making a choice as to whether to see the movie. Following the movie's release, however many people start to watch the movie, so that potential moviegoers can receive more precise information from their peers. The uncertainty about movie quality that can be resolved over time involves a Bayesian method which depends upon the number of previous viewers and their ratings reported over the Internet. This paper separately consider these two social learning effects and, in particular, focuses on the second type of social learning.

Critics' Reviews

Critics' reviews are generally regarded as a good source of information about movie quality. Although they do not perfectly reflect consumers' perceived qualities in future, they are highly correlated with consumer attitudes about a movie's quality. In our data, the Pearson's correlation coefficient between Metascores and IMDb ratings is 0.7466, and the Spearman's rank correlation coefficient between them is 0.7479.

Critics' reviews are influential, particularly, around the time of a movie's release. According to Moretti (2008), 85% of newspaper reviews are published on or before the date of the opening. Because critics reviews are available earlier than other social learning factors, so they play a key role in constructing early belief about movie quality and hence consumers' movie-watching choices. In that sense, critics ratings can indicate the movie's success given other movie characteristics, which, in turn, can influence studios' advertising decisions.

Timing of Advertising

To justify the assumption that studios' advertising decision can be simplified as a single time decision for the total advertising budget, we have to explain how advertising expenditures are distributed over time. Once a movie is produced, a studio starts marketing it. The movie trailers are screened at theaters and its internet web-page is often launched providing the information about the director, casts, plot and so on. As the release week gets closer, the studio advertises the movie on numerous media outlets including television, newspapers, magazines, and posters. All expenditures spent before the release accounts for more than 50% of total adverting spending. During the opening week, to increase audience awareness of the movie, the studio advertises extensively through the media. Indeed, about 25% of the total advertising spending occurs during the opening week. For the movies in our data set, approximately 78% of all advertising expenditure is distributed in the market prior to the first week of the movie's release.

Since most occurs prior to are immediately following a movie's release, it is not unduly unrealistic to consider total advertising expenditures rather than weekly advertising expenditures. One issue that must be considered, however, is the effects of movie critics reviews on advertising decisions. Critics' reviews become available just prior to the film's release date. By that time, only 50% of total advertising expenditures have been allotted. Thus, the other 50% of advertising spending, can be influenced by good or bad reviews. Throughout this paper, we account for this relationship between advertising and critics reviews by use of an instrumental variable. The most plausible instrumented variable for this measure is production budget of a film. The presence of famous directors, actors or actresses are positively related to movie quality, which, in turn, can increase the production budget of a film. Production costs, however, are sunk costs so that neither advertising nor critics influence production costs. Since the total production budget then is correlated with both movie quality and advertising spending but not influenced by either of them, it serves as a good instrumental variable for use in our estimation.

4 Preliminary Discussions of Social Learning and Advertising

4.1 Multiplier Effect of Social Learning on Advertising

If social learning has a positive multiplier effect on movie advertising, movie distributors should spend more on movies which are expected to acquire positive word-of-mouth recommendations and less on those expected to get bad word-of-mouth referrals. On the other hand, if social learning has a negative multiplier effect on advertising, studios should increase their advertising budgets for low-rated films to compensate for bad reviews and (using the same logic) to reduce them for high-rated films. Thus, the studio's profit-maximization and optimal-advertising decisions can be influenced by whether the multiplier effect of social learning for movie advertising expenditures is positive or negative.

The cumulative market share (or the cumulative box office revenue) can be considered as evidence of social learning, since the popularity of a movie can be interpreted as a signal of good quality. However, to consider the direct effect of social learning and to explore the multiplier effect, we consider variables from which movie-goers learn about movie quality, such as critics' reviews and viewers' ratings.

First, critics' ratings are regarded as reliable information about a movie's quality before its release. For critics' ratings, we use the "Metascores" available through *Metacritic* (www.metacritic.com). *Metacritic* is a website which provides a vast amount of critical information about entertainment goods, including films. Metascores combine individual critic scores into a composite grade for each movie, based on a scale from 0 to 100. The mean of the average Metascores for all movies in our dataset is 50.0 and the mean of the SDs for all Metascores is 14.5. Other variables that may affect individuals' perceptions of movie quality are the Internet ratings from previous movie-goers. Internet ratings can be considered as quantified perceived qualities. Here, we employ user ratings collected from the *Internet Movie Database* (IMDb). The IMDb is a popular website in which users receive information about films and share their opinions about movies by rating them on a scale from 0 to 10. The mean of the IMDb ratings of movies in our dataset is 6.28, and the mean of the SDs of those ratings is 2.15. Metascores and IMDb ratings are highly correlated but not the same: the Pearson's correlation coefficient between them is 0.7466, and the Spearman's rank correlation coefficient between them is 0.7479. To avoid the problem from which Metascores and IMDb ratings use different scales, we transform the Metascores to the same scale as the IMDb ratings.

The following regression results accounts for the effects of these two social learning components, Metascores and IMDb ratings, as well as the effect of total advertising expenditure on the log of total box-office revenue:

$$\ln(bo_j) = a_0 + a_1 \cdot ad_j + a_2 \cdot cr_j + a_3 \cdot ir_j + a_4 \cdot ad_j \cdot cr_j + a_5 \cdot ad_j \cdot ir_j + e_j,$$

where bo_j is the total box-office gross of movie j in millions of dollars, ad_j is the total advertising expenditure for movie j in 10 millions of dollars, cr_j is the average Metascore for movie j transformed to the same scale as the IMDb ratings and ar_j is the average IMDb rating of movie j, e_j is the disturbance of movie j. Here, the endogeneity problem discussed above can be easily seen. As we explain in Section 3.2, about 50% of total advertising spending is decided upon during or after a movie's release week, so that ad_j can be correlated with bo_j. Thus, we consider the total production budget as an instrumental variable for this estimation.

Table 5 provides the IV regression results. The results from Columns (1) through (5) show the effects of the independent variables on the log of total box-office revenues based on all movies in our dataset when some variables are considered regressors and others are not. Column (1) shows the marginal effect of advertising spending on the box-office revenue when no social learning factors are considered. Column (2) shows the marginal effects of both advertising expenditures and Metascores. As can be seen if the effects of critics' ratings are considered, the marginal effect of advertising spending decreases a little. Column (3) shows not only the marginal effects of advertising spending and Metascores but also their joint effect. Column (3) suggests two facts worth noting. First, when the joint effects are considered, the parameter estimate for critic rating becomes negative, but is not statistically significant. Second, the parameter estimate for each variable decreases relative to that in Column (2). Second, the parameter estimate for the joint term is positive and statistically significant. The results in Column (3) then implies that critics' ratings have a positive multiplier effect on advertising spending. From Column (3), the following equation can be obtained:

$$\begin{aligned}\ln(bo_j) &= \underset{(0.843)}{2.291} + \underset{(0.119)}{0.431} \cdot ad_j - \underset{(0.020)}{0.016} \cdot cr_j + \underset{(0.003)}{0.045} \cdot ad_j \cdot cr_j,\end{aligned}$$

where the numbers in parentheses are the standard errors. From this equation, the partial effect of advertising conditional on the critics rating can be computed as follows:

$$\left.\frac{\partial \ln(bo_j)}{\partial ad_j}\right|_{cr_j} = 0.431 + 0.045 \cdot cr_j.$$

This implies that when a higher critic review score is given, the effectiveness of advertising increases. This, in turn, supports the notion that social learning has a positive multiplier effect on advertising. Column (4) shows the marginal effects of advertising spending, Metascores and IMDb ratings. The marginal effect of Metascores becomes lower relative to Column (2), implying that some part of social learning is explained by critics' ratings and part of the remainder is explained by word-of-mouth effects. Column (5) shows both the marginal effects of these three variables and the joint effects between advertising and each of the social learning factors. Similar to Column (3), Column (5) provides the following equation to show the partial effect of advertising conditional on social learning:

$$\left.\frac{\partial \ln(bo_j)}{\partial ad_j}\right|_{cr_j, ir_j} = 0.382 - 0.009 \cdot cr_j + 0.006 \cdot ir_j$$

The partial effect here is always positive, because cr_j is less than 10. This implies that word-of-mouth referrals have an important role in increasing the effectiveness of movie advertising.

Next, in order to see whether this positive multiplier effect is strong for either good or bad movies, we divide the movies in our dataset into three groups, with respect to the mean of Metascores for each film. Group 1 consists of films with high Metascores; Group 2 is comprised of films with medium Metascores; and Group 3 contains those films with low Metascores. Columns (6), (7) and (8) in Table 5 show the marginal effects of advertising spending, Metascores and IMDb ratings on the box-office revenue in each group. Interestingly, the marginal effect of advertising spending decreases from Group 1 to 3. That is to say, the effectiveness of movie advertising is stronger for good movies and relatively weaker for bad movies. Furthermore, for high-quality movies, the marginal effect of IMDb ratings is negative, while for average-quality movies, the marginal effects of Metacritics and IMDb ratings are negative. These results may imply that without a sufficiently large advertising campaign, social learning for good or average movies cannot bring in large box-office revenue, and the finding is consistent with a positive social learning multiplier effect on advertising.

4.2 Social Learning's Effect on Box Office Gross

Next, we consider how the social learning process relates to weekly decay patterns in movie demand, focusing on movies that were initially widely released. The weekly box-office revenue is influenced by many observed and unobserved movie characteristics. Particularly, the box-office gross for a movie's release week is affected by all other factors, excluding word-of-mouth effects. If good or bad word-of-mouth reports influence box-office revenue, the ratio of the weekly box-office gross to week 1 box-office revenue changes can be highly correlated with social learning factors, (*i.e.*, Metascores and IMDb ratings). That is, if both Metascores and IMDb ratings are large, then the relative ratios can be high, which means that social learning factors increase movie consumption. The following regression captures how two social learning components can influence the ratio of box-office revenue for each subsequent week to that of the movie's initial release week:

$$\ln\left(\frac{bo_{jt}}{bo_{j,t=1}}\right) = b_0 + b_1 \cdot ad_j + b_2 \cdot cr_j + b_3 \cdot ir_j + e_{jt},$$

where bo_{jt} is the box-office revenue of movie j in millions of dollars in week t following the initial time of release and $BO_{j,t=1}$ is the box-office revenue of movie j in millions of dollars during the release week. e_{jt} is the disturbance of movie j in week t. As opposed to the situation described in Section 4.1, the endogeneity problem here can be avoided. Here, the weekly box office revenues are controlled by the first week's box office gross, so that most advertising spending is independent of weekly box office performances controlled by the opening week's box office revenue. Thus, OLS regression estimation provides a suitable means for addressing this question.

The OLS regression results for Week 2 through Week 5 are given in Table 6. From Weeks 2 to 5, the effect of the constant term decreases, while those of two social learning terms and advertising spending increase. The value of R^2 goes up as well. The results imply three things. First, the marginal effect of the constant term decreases because the effect of the movie's unobserved factors declines relatively over time. Second, social learning plays a more important role in promoting movie consumption as time goes by. Finally, a large volume of advertising spending continuously influences movie demand with social learning factors, which also can be related to the multiplier effect of social learning on advertising. In particular, these results indicate that good word-of-mouth reports can smooth out the steepness of movie decay patterns, yielding additional box-office sales. These OLS estimates then suggest that good word-of-mouth referrals increase the movie's demand, while bad word-of-mouth referrals lower it.

5 Model

In this section, we develop a structural model for movie demand by heterogeneous households and optimal advertising decisions for motion picture distributors. The model assumes that each movie has sufficient monopolistic power so there is no direct competition between movies. However, the relationship among movies is examined indirectly here in the form of a relative evaluation of movie quality. Studios make a one-time decision for advertising expenditure, but they fully consider the future revenue stream of movie demand so that they are by no means myopic. Consumers, however, make a repeated myopic movie-watching decision until they either see the film or the film is no longer showing. Furthermore, they watch a movie at most once.

5.1 Household Demand for Movies

We assume that the duration of a movie's theatrical release is both given and known to the consumer. Week 1 stands for the opening week of movie j in theaters and let D_j denote the total duration in weeks of movie j's release.

In week t, $1 \leq t \leq D_j$, household i has two options: to see movie j or not. If the household watches the film, its payoff can be written as:

$$u_{ijt} = \sum_k x_{jkt} \left(\beta_k + \gamma_k v_{ik}\right) - \alpha_i \ln\left(p_{ij}\right) + \sum_k \rho_k T_{jkt} + \xi_{jt} + \left(\varepsilon_{ijt} - \varepsilon_{i0t}\right), \tag{1}$$

where:

- x_{jkt} is the observed movie characteristic of movie j in week t,
- β_k is the *common* taste parameter associated with movie characteristic k,
- γ_k is the parameter measuring the heterogeneity in tastes,
- v_{ik} is the household attribute,
- α_i is the marginal utility of household i's income,
- p_{ij} is the *total* price of movie tickets that household i has to pay for watching movie j, taking into account the member of adults and children in the household and whether or not the movie is suitable for children,
- T_{jkt} is the seasonal dummy variable in week t,
- ρ_k is the parameter associated with time-specific characteristic k,
- ξ_{jt} is the unobserved characteristic of movie j in week t that is not explained by the observed movie characteristics, the price or the seasonal dummy variables, and

- ε_{ijt} and ε_{i0t} are, respectively, the idiosyncratic error terms when the household watches movie j in week t and when it does not.

In our example, x_{jkt} is a vector of the movie characteristics (e.g., genre, weekly adjusted effect of initial advertising expenditure, cumulative market share of previous viewers, weekly common belief of movie quality, Academy Award nomination and Academy Award receipt). Among more than ten genres, we consider two of the most popular classes of films: family films and action movies. Following Moul (2006), family films are those rated G or PG and are therefore suitable for young children; action movies are those classified as "Action" or "Adventure" by *Variety* (www.variety.com).

Some movie characteristics such as the effect of advertising change over time. Because most advertising investments occur prior to or upon a movie's release, we consider the total advertising spending on the theatrical market as an indicator of advertising's effect for the release week. We assume that exposure to advertising potentially decreases over time and the rate of advertising depreciation is given by η (which is estimated). Hence, the weekly adjusted advertising effect can be written as:

$$(1-\eta)^{t-1} \cdot ad_j,$$

where ad_j represents the total advertising expenditure on movie j. These adjusted advertising effects are one of the observed movie characteristics. A film's overall popularity can be interpreted as circumstantial evidence of its quality, and such popularity can promote additional movie consumption. To reflect this, the cumulative market share of previous viewers is also included as one of the observed characteristics.[9] In the first week of a movie's release, critics' ratings are the most reliable information about the movie's quality. After the movie is shown in theaters, word-of-mouth recommendations from viewers become available, and such recommendations serve to continuously update the average reputation about the movie's quality over time. Information about movie quality, which is initially based on critics' ratings, is thus updated in a Bayesian method that depends on the number of previous film-goers and their Internet ratings. The last two observed variables

[9]This implies that the viewing decisions of others influence my decisions and my decisions, in turn, impact others. Such a relationship requires that the equilibrium of consumers' decisions follows a Nash equilibrium. To reflect this equilibrium concept, we use the cumulative market shares simulated from the estimation process, not those observed from the actual data.

are a film's Academy Award nominations and receipts. The Academy Award is the premier film award in the U.S. with movies receiving Oscars or nominations generally regarded as well-made films and that reputation tends to further raise movie attendance.

The random variate, v_{ik}, reflects the heterogeneity in how a household values the observed movie characteristic. Following Berry, Levinsohn and Pakes (1995, 2004), v_{ik} is treated as the household's attitude. The number of children influences movie-watching decision of a household in that households with children are more likely to go to the theaters for family movies than for movies of other genres. In fact, the age of the youngest child appears to be an important factor; if elderly children want to see a family movie but the youngest child is too young to go to the theater, the household may be required to hire a babysitter for the outing, compounding its cost. Because it is not recommended to bring children younger than 6 years old to the theater, the age of 6 is the critical age for influencing a household's movie-going behavior. So, v_{ik} can be rewritten as:

$$v_{ik} = \begin{cases} v_{ik}^u + \zeta_6 \cdot \mathcal{I}\left(age < 6\right) + \zeta_{17} \cdot \mathcal{I}\left(6 \leq age \leq 17\right), & \text{if } k = \text{family genre,} \\ v_{ik}^u, & \text{otherwise,} \end{cases}$$

where v_{ik}^u is the unobserved household attitude, $\mathcal{I}$ is the indicator function and age refers the age of the youngest child in the household.

We allow the impact of movie price on utility, α_i, to vary according to income groups:

$$\alpha_i = \begin{cases} \alpha_1 & \text{if } y_i < \overline{y}_1 \\ \alpha_2 & \text{if } \overline{y}_1 \leq y_i < \overline{y}_2 \\ \alpha_3 & \text{if } y_i \geq \overline{y}_2, \end{cases}$$

where y_i represents the individual household's income and $\overline{y}_1$ and $\overline{y}_2$ are the cutoff income levels which serve to divide the U.S. population into 3 groups for this study. p_{ij} is the ticket price. Although a ticket price does not change week to week, the total ticket cost that each household faces varies with respect to the family's composition as well as the film's MPAA rating.[10] The MPAA rating system decides which movies are considered appropriate for children and/or adolescents. Thus, which family members can watch a movie depends on its MPAA rating, and in that sense, the ticket price is indexed by both household i and movie j. For these two terms, α_i and $\ln\left(p_{ij}\right)$, we

[10] A ticket price increases by roughly 3.7% each year. However, if we consider the overall inflation level and adjust the price using Consumer Price Index (CPI) data, the adjusted ticket price does not change significantly. So, throughout the model, we assume that the ticket price for a single person does not change over time.

need three attributes for each household: income, family size and age distribution of the household. We denote v_i^o by the vector of these observed household attributes. They are available from the IPUMS, which provides census data with extensive information using a large random sample of U.S. households.

T_{jkt} is the weekly time-specific variable, such as a holiday-weekend dummy and a summer-season dummy. Although T_{jkt} is not directly related to movie characteristics, it can make a difference in weekly box office figures. During holiday weekends, for example, people have more time for leisure, so they are more likely to go to the theaters, raising the weekly box office revenue. Hence, we include a dummy variable indicating whether a weekend falls immediately prior to one of six holidays: President's Day, Memorial Day, Independence Day, Labor Day, Thanksgiving and Christmas. Additionally, since movie consumption tends to increase during the summer, we include a variable indicating whether a weekend lies between Memorial Day and Labor Day.

ξ_{jt} represents the unmeasured aspect of the movie and is not restricted to be constant over time. ξ_{jt} can be decomposed into a permanent component, a time since release effect, and a random iid error. Thus, we have

$$\xi_{jt} = \xi_j + \xi_t + \Delta\xi_{jt}, \tag{2}$$

where ξ_j is the movie-specific unobserved factor, ξ_t is the time trend and $\Delta\xi_{jt}$ is the remainder. Although weekly movie consumption exhibits a general decay pattern, individual movie consumption patterns are highly heterogeneous. To describe this heterogeneity within the unobserved characteristics, we include the fixed effect, ξ_j. Because ξ_j has influence on movie j regardless of the number of weeks elapsed, ξ_j varies among movies, but is constant over time for each movie considered. There may also exist unexplained factors influencing the decay pattern which depend on the number of weeks elapsed but are independent of the movies themselves. To describe this time-specific unobserved characteristic, we consider the time-specific term, ξ_t. ξ_t changes over time, but it is the same for all movies in the same week. Finally, $\Delta\xi_{jt}$ is the remainder that is not explained by either ξ_j or ξ_t, the econometric error term. ε_{i0t} is included in u_{ijt} rather than u_{i0t} for the purpose of normalizing the utility of the outside good (from not watching movie j at a theater in week t) to be 0. Thus, we include the term $-\varepsilon_{i0t}$ in u_{ijt}.

We assume that each household watches a movie no more than once. To reduce computational complexity, we also assume that a given household makes a decision based only on the available information in the present week without considering the future payoff, but they repeatedly choose whether to go to the movie if they have not seen that movie before. Let I_{ijt} be the information set about movie j given to household i in week t. I_{ijt} can be written as:

$$I_{ijt} = (x_{jt}, p_{ij}, T_{jt}, \xi_{jt}), \ 1 \leq t \leq D_j.$$

Next, let d_{ijt} be household i's movie-watching decision in week t. d_{ijt} is denoted as 1, when the household goes to movie j, and 0 when it does not. For household i then with the history $\{d_{ij\tau}\}_{\tau=1}^{t-1} = 0$, the decision problem in week t is given as:

$$V(I_{ijt}| \{d_{ij\tau}\}_{\tau=1}^{t-1} = 0) = \max \{u_{ijt}, 0\}, \ 1 \leq t \leq D_j.$$

5.2 Information Update about Movie Quality

To see the process in which the information about movie quality is generated and becomes more accurate week by week, we assume that the common belief of movie quality is updated by means of a Bayesian process and consider its distribution.

For week $t = 1$, it is assumed that no person has watched movie j, so the best guess for a movie's quality is based upon the distribution of critics' ratings. Let q_{jt} denote the common belief about the quality of movie j in period t where q_{jt} is one of the observed movie characteristics (*i.e.,* $q_{jt} \in (x_{jkt})$ for $1 \leq t \leq D_j$). We can write the initial prior for movie j as:

$$q_{j,t=1} \sim N(m_{cr_j}, \sigma^2_{cr_j}), \tag{3}$$

where m_{cr_j} is the average of critics' ratings for movie j and $\sigma^2_{cr_j}$ is their variance. In week 1, households get noisy signals about the quality drawn from this initial prior distribution. If $\sigma^2_{cr_j}$ is large, people can receive information that the movie quality is poor even for good movies. However, if $\sigma^2_{cr_j}$ is small, the uncertainty is small, so people can receive more accurate information about the movie's quality. As time passes, other people watch the movie and share their opinions, and however the more people watch, the more this information is updated. Due to taste heterogeneity, their quality evaluations are not necessarily identical. Let n_{jt} denote the total number of people who have seen movie j up to period $t-1$. To distinguish the perceived qualities from the common belief for quality, we denote $\phi_{\iota,j}$ as the perceived quality of the ι-th movie-goer among n_{jt} (*i.e.,* $1 \leq \iota \leq n_{jt}$). $\phi_{\iota,j}$ has a true distribution that is *ex-ante* unknown and denoted by $N(m_{\phi_j}, \sigma^2_{\phi_j})$. m_{ϕ_j} is the average of perceived qualities for movie j, and $\sigma^2_{\phi_j}$ is their variance. [11] According to the Bayesian rule, a potential consumer receives updated information about the movie that follows a posterior distribution given as:

$$q_{jt} \sim N(m_{jt}, \sigma^2_{jt}), \tag{4}$$

where

$$m_{jt} = \left(\frac{1}{\sigma^2_{cr_j}} + \frac{w}{\sigma^2_{\phi_j}} n_{jt} \right)^{-1} \left(\frac{m_{cr_j}}{\sigma^2_{cr_j}} + \frac{w}{\sigma^2_{\phi_j}} \sum_{\iota=1}^{n_{jt}} \phi_{\iota,j} \right) \text{ and } \sigma^2_{jt} = \left(\frac{1}{\sigma^2_{cr_j}} + \frac{w}{\sigma^2_{\phi_j}} n_{jt} \right)^{-1}, \tag{5}$$

where w is the velocity of the information update. In reality, not all people watching the movie share their opinions, which slows the update to the common belief. Therefore, we assume that w is less than or equal to 1, *i.e.,* $w \in [0, 1]$. It is difficult to estimate w. Instead, we fix w at a reasonable number that makes the update velocity neither too fast nor too slow.[12] The posterior distribution requires that the equilibrium of consumers' decisions follows a Nash equilibrium. So, we use the number of movie-watchers simulated from the estimation process, not observed from the actual data. Similar to week 1, households get noisy signals about the quality drawn from this posterior distribution. If σ^2_{jt} is large, the uncertainty about the quality is large, and, if $\sigma^2_{cr_j}$ is small, the uncertainty is small. Generally, as time passes, more people watch the movie and n_{jt} increases. The increase of n_{jt} reduces σ^2_{jt}. So, the uncertainty is resolved over time and people finally get very accurate information about the movie quality. Actually, n_{jt} is used not only for updating the posterior distribution but also for a cumulative market share, another variable among the observed movie characteristics. Through the social learning process, an increase in n_{jt} reduces the variance of the posterior distribution. That is, n_{jt} reduces the noise of the quality indicator, q_{jt} over time, but does not directly increase or decrease q_{jt}. On the other hand, n_{jt} directly affects

[11] For estimation, the distribution of IMDb ratings is used for the distribution of $\phi_{\iota,j}$. Therefore, m_{ϕ_j} is replaced with ir_j, the mean of IMDb ratings of movie j, and $\sigma^2_{\phi_j}$ is substituted with the SD of IMDb ratings of movie j.

[12] In this paper, w is fixed as 1.0e-007, with which movie consumers can get quite accurate information about movie quality four weeks later. This is consistent with the OLS results in Section 4.

the cumulative market share because the cumulative market share is n_{jt} divided by the population. So, n_{jt} continuously increases the variable of the cumulative market share variable.

To derive the above probability distributions of average qualities, we use both the Metascores and the IMDb user ratings. Metascores and IMDb ratings use different scales, so we transform the Metascores to the same scale as the IMDb ratings. Then, we construct the prior and the posterior distributions for the common belief of movie quality.

5.3 Estimation of Demand Model

Let p_j be the ticket price that the *average* household faces when watching movie j and $\overline{\alpha}$ be its parameter for the average household. In the theatrical market, the choice-specific constant term, δ_{jt}, can be given as:

$$\delta_{jt} = \sum_k x_{jkt}\beta_k - \overline{\alpha}\ln(p_j) + \sum_k \rho_k T_{jkt} + \xi_{jt}. \tag{6}$$

By substituting (6) with (1), household i's utility of going to movie j can be transformed as:

$$u_{ijt} = \delta_{jt} + \sum_k x_{jkt}\gamma_k v_{ik}^u - (\alpha_i \ln(p_{ij}) - \overline{\alpha}\ln(p_j)) + (\varepsilon_{ijt} - \varepsilon_{i0t}),$$

or

$$u_{ijt} = \delta_{jt} + \sum_k x_{jkt}\gamma_k v_{ik}^u - \alpha_i (\ln(p_{ij}) - \ln(p_j)) - (\alpha_i - \overline{\alpha})\ln(p_j) + (\varepsilon_{ijt} - \varepsilon_{i0t}). \tag{7}$$

Because q_{jt} is one of the observed movie characteristics, *i.e.*, $q_{jt} \in (x_{jkt})$, q_{jt} enters u_{ijt} in (7) through the constant term, δ_{jt}, and the heterogeneous term, $\sum_k x_{jkt}\gamma_k v_{ik}^u$.

To obtain simple expressions for choice probabilities conditional on I_{ijt}, we assume that ε_{ijt} and ε_{i0t} are I.I.D. extreme valued random variates. In week 1, each household receives a noisy signal about q_{jt} that is drawn from the prior distribution. Similarly, in week t, the household gets a noisy signal about q_{jt} drawn from the posterior distribution. In the structural model, the probability that movie j is watched by potential consumer i in period t, $1 \leq t \leq D_j$, given the information set I_{ijt} is:

$$\Pr\left\{d_{ijt} = 1 | \{d_{ij\tau}\}_{\tau=1}^{t-1} = 0\right\} = f_{ijt} \cdot \prod_{t'=1}^{t-1}(1 - f_{ijt'}),$$

where

$$f_{ijt} = \Pr\{d_{ijt} = 1\} = \int_q \frac{\exp(u_{ijt})}{1 + \exp(u_{ijt})} g(q) dq,$$

where f_{ijt} is the probability to see movie j unconditional on the history of movie consumption, and q means q_{jt} and $g(\cdot)$ is the density of q_{jt}. f_{ijt} is computed by the integration on $g(\cdot)$, the distribution of the random variable, q_{jt}.

Denote V_i by the vector of all household attributes including both v_i and v_i^o, (*i.e.*, $V_i = (v_i, v_i^o)$), and $\mathcal{P}_V$ by its distribution in the population. Let s_{jt} be the proportion of *individuals* going to movie j in week t. From (6) and (7), s_{jt} can be written as:

$$s_{jt}(\delta(\theta), \theta | x, \mathcal{P}_V) = \frac{H}{N} \int hh_{ij} \Pr\left\{d_{ijt} = 1 | \{d_{ij\tau}\}_{\tau=1}^{t-1} = 0\right\} dV, \tag{8}$$

where $\theta = (\beta, \alpha, \gamma, \rho, \mu)$. N is the total population in the U.S., H the total number of households in the U.S. and hh_{ij} the number of potential movie consumers of movie j in household i.

Unfortunately, neither $s_{jt}(\delta(\theta), \theta | x, \mathcal{P}_V)$ nor f_{ijt} has a closed form solution. Consequently, we follow Berry, Levinsohn and Pakes (2004) and use simulation to approximate their values.

5.4 Optimal Advertising Decision of Motion Picture Distributors

As noted in the introduction, we focus on the studio's profit maximization problem once a movie is produced. After a film is made, the production budget becomes a sunk cost, so it is assumed to no longer be a decision variable for the movie studio. After the movie is made, but before it is released to the market, the distributor's main concern is how effectively it can advertise the movie and maximize its profit. In this paper, studios make a one-time decision for advertising expenditure, but they fully consider the expected future revenue stream of the movie. To construct the studio's profit maximization problem, we look first at the studio's net profit for each week. We denote the week 0 as the time period during which the movie has been produced but not yet released.

In week t, $0 \le t \le D_j$, the studio's net profit for movie j is given as:

$$\pi_{jt}(ad_j) = \begin{cases} -ad_j, & \text{if } t = 0, \\ \psi(p \cdot s_{jt} \cdot N - c \cdot sc_{jt}), & \text{if } 1 \le t \le D_j, \end{cases}$$

where $\pi_{jt}(\cdot)$ is the net profit in week t, ad_j is the total advertising spending, ψ is the distributor's share, p is the price of one ticket, c is the operation cost per screen (the so-called "nut") and sc_{jt} is the number of screens. In week 0, the studio must spend money on advertising, but the movie has yet to generate a revenue stream. So, its net profit comes from only the cost side. After a movie's release, the box office revenue is collected by theaters. Before the revenues are shared, the "nut" is given to theaters. Then, the net revenue (excluding the operational cost) is shared between the studio and theaters. In other words, the studio's net profit in week t is some proportion of the total box office revenue, net of the nut. The total box office revenue is computed by the price of one ticket and the total ticket sales. In turn, the total ticket sales are equal to the total population multiplied by the market share. The weekly market share, s_{jt}, comes from the market demand based on households' movie-watching decisions in (8). The weekly market share is a function of the level of advertising, ad_j. Also, the number of engaged screens in each week is just a deterministic function of the market demand in that week. Thus, the number of theaters can be a function of the market share. Since the market share depends on the total advertising level, the number of screens will be given as:

$$sc_{jt} \equiv \Lambda\left(s_{jt}\right) \equiv \Lambda'\left(ad_j | j, t\right).$$

Therefore, the weekly net profit can be expressed as a function of the total advertising expenditure. The potential difficulty in computing the net profit stream is that the market shares rely on the social learning process, but the distribution of perceived qualities is *ex-ante* unknown. To compute the net profit, the studio must forecast the distribution of perceived qualities. We assume that the studios use the following equations to forecast the mean and variance of the quality distribution:

$$E\left[m_{\phi_j}\right] = \mu_0 + \mu_1 cr_j + \mu_2 (cr_j)^2 + \mu_3 pc_j$$

and

$$E\left[\sigma_{\phi_j}\right] = \frac{1}{J}\sum_{j'} \sigma_{\phi_{j'}},$$

where pc_j is its production cost and cr_j is the initial critic rating.

Next, consider the studio's profit maximization problem. Because the stream of weekly net profits depends on the advertising decision, the studio's optimization problem is given as:

$$\max_{ad_j} \pi_j = \sum_{t=0}^{D_j} \frac{1}{(1-r)^t} \pi_{jt}\left(ad_j\right),$$

where r is the discount rate. The first order condition is given as:

$$\sum_{t=0}^{D_j} \frac{1}{(1-r)^t} \frac{d\left(\pi_{jt}\left(ad_j\right)\right)}{d\ ad_j} = 0.$$

From this condition, the optimal advertising level, ad_j^*, is determined.

6 Estimation Strategy

We estimate the model using the SMM estimation. To evaluate the out-of-sample fit by the model, we randomly leave out 36 movies from the dataset and assess how well the estimated model forecasts the decay patterns for those movies in 7.2.

6.1 Moments

Here, we describe how to compute the moments that go into the SMM estimation algorithm. We follow a similar identification strategy to that of Berry, Levinsohn and Pakes (2004) and Petrin (2002). We use the two sets of moments. The first set of moments consists of both studio-side and demand-side disturbances, which can be obtained by solving the equilibrium of studios' optimal advertising strategies and consumers' movie demand. To get the demand-side disturbance ($\Delta\xi_{jt}$), we first have to obtain the choice-specific constant term (δ_{jt}). However, it is not simple to get δ_{jt} because of the heterogeneity among households and the myopic but repeated household consumption style in this paper. So, to get δ_{jt}, we follow Berry, Levinsohn and Pakes (1995) and Gowrisankaran and Rysman (2006) and use the contraction mapping for δ_{jt} with aggregate market data. The second set of moments can be constructed by finding the relationship between some household attributes and movie characteristics with consumer-level data. our consumer-level data is the *Simmons* NCS fall data set. It shows how consumption of movies varies across aggregate groups of households, but it does not provide detailed household level information, such as each household's movie consumption behaviors. Thus, we include two relationships in the moment conditions: (i) the consumption variations of the three income groups and (ii) the ratio of the movie-going probability in each group categorized by the age of the youngest child to that in the total population.

The first set of moments and δ_{jt} relates to the studio-related disturbance (ω_j) and market-level demand disturbance ($\Delta\xi_{jt}$). Unlike other papers, in our model prices are treated as exogenous variables. The price of a movie ticket is fixed regardless of movies. In other industries (such as the car industry), prices reflect unobserved product characteristics like brand effects. However, in the movie industry, the ticket price for the movie made by a famous director or starring a popular actor is the same as that of other movies. So, the prices themselves are not correlated with the disturbances. However, the advertising expenditure can be correlated with ω_j and $\Delta\xi_{jt}$, because the advertising decision is made based on the expected box office revenues and movie demand is influenced by advertising spending. So, ω_j and $\Delta\xi_{jt}$ are assumed to be uncorrelated with all observed variables except advertising factors on the demand-side and those on the studio-side. We denote z_{jt} by the vector of variables including p_j, T_{jt}, $(cr_j)^2$, pc_j and all x_{jkt} except weekly advertising effects. Then, the first set of moments can be derived from:

$$E\left[\Delta\xi_{jt}\left(\theta_0\right)|z_{jt}\right] = E\left[\omega_j\left(\theta_0\right)|z_{jt}\right] = 0.$$

ω_j and $\Delta\xi_{jt}$ can be correlated with each other, but both of them are orthogonal to z_{jt}. Thus, the first set of moments is given as:

$$G_1\left(\theta_0\right) = E\left[H_j\left(z_{jt}\right)T_j\left(z_{jt}\right)\begin{pmatrix}\Delta\xi_{jt}\left(\theta_0\right)\\ \omega_j\left(\theta_0\right)\end{pmatrix}\right],$$

where $H_j\left(z_{jt}\right)$ is the matrix of instruments, z_{jt}, and $T_j\left(z_{jt}\right)$ is the matrix that standardizes the disturbances and satisfies:[13]

$$T_j\left(z_{jt}\right)'T_j\left(z_{jt}\right) = \left(E\left[\left(\Delta\xi_{jt},\omega_j\right)'\left(\Delta\xi_{jt},\omega_j\right)|z_{jt}\right]\right)^{-1}.$$

To compute $\Delta\xi_{jt}\left(\theta\right)$, we exploit the fixed point equation in accordance with Berry, Levinsohn and Pakes (1995) and Gowrisankaran and Rysman (2006). According to (6), δ_{jt} can be written as $\delta_{jt} = \sum_k x_{jkt}\beta_k - \overline{\alpha}\ln\left(p_j\right) + \xi_{jt}$. Then, using the contraction mapping theorem, we obtain the fixed point of δ_{jt}:

$$\delta_{jt}^{new} = \delta_{jt}^{old} + \left[\ln\left(s_{jt}\right) - \ln\left(\hat{s}_{jt}\left(\delta_{jt}^{old},\theta\right)\right)\right], \tag{9}$$

[13] In this paper, $H_j\left(z_{jt}\right)$ is the 12 x 2 matrix and $T_j\left(z_{jt}\right)$ is the 2 x 2 matrix.

where s_{jt} is obtained from the data and $\hat{s}_{jt}\left(\delta_{jt}^{old},\theta\right)$ is calculated from (8).[14] s_{jt} is the ratio of the number of people watching movie j in week t to that of all people in the U.S. For the computation of s_{jt}, we need the weekly ticket sale information. However, the available information is expressed as total box office revenues, not total ticket sales. Thus, we calculate the number of total ticket sales using weekly box office revenues and average ticket prices. We generate the weekly average ticket prices as linearly transforming the yearly average ticket prices, divide weekly box office revenues by weekly average ticket prices, compute the weekly ticket sales, and obtain s_{jt}. Given θ, ξ_{jt} can be computed from (9). In turn, using the equation (2), $\Delta\xi_{jt}$ can be obtained.

The second set of moments matches (i) the model's prediction of the movie-going probability (at least once) conditional on household income levels with that of the NCS 2002 fall data set and (ii) the model's estimate for the ratio of the movie-watching probability in each group categorized by the youngest child's age to the movie-going probability in the population with that of the NCS 2002 fall set. Movies used for generating these moments do not completely overlap with those we actually focus on. However, with the consistency assumption of household attributes, we can use this data set. We denote movie j' by the j'-th movie in the NCS 2002 fall set and F by the set of family movies. For three income groups, the moments are given as:

$$\begin{aligned}
s_{j'}^{NCS}\left(y<\overline{y}_1\right)-\int_{y<\overline{y}_1}\left(1-\Pr\left\{\left\{d_{ij't}\right\}_{t=1}^{D_{j'}}=0\right\}\right)dV &= 0,\\
s_{j'}^{NCS}\left(\overline{y}_1\leq y<\overline{y}_2\right)-\int_{\overline{y}_1\leq y<\overline{y}_2}\left(1-\Pr\left\{\left\{d_{ij't}\right\}_{t=1}^{D_{j'}}=0\right\}\right)dV &= 0,\\
s_{j'}^{NCS}\left(y\geq\overline{y}_2\right)-\int_{y\geq\overline{y}_2}\left(1-\Pr\left\{\left\{d_{ij't}\right\}_{t=1}^{D_{j'}}=0\right\}\right)dV &= 0,
\end{aligned}\tag{10}$$

where $s_{j'}^{NCS}\left(y<\overline{y}_1\right)$ is the proportion of households with income less than $\overline{y}_1$ going to movie j' in any week, $s_{j'}^{NCS}\left(\overline{y}_1\leq y<\overline{y}_2\right)$ is the proportion of households with income between $\overline{y}_1$ and $\overline{y}_2$ going to movie j' in any week, and $s_{j'}^{NCS}\left(y\geq\overline{y}_2\right)$ is the proportion of households with income greater than or equal to $\overline{y}_2$ going to movie j' in any week. Also, for the other three groups categorized by

[14]Berry(1994) provides the conditions under which this function is a contraction mapping, guaranteeing that $\hat{s}_{jt}\left(\delta_{jt}^{old},\theta\right)$ is invertible in the vector of δ's. Although our model doesn't satisfy all of Barry's conditions, we haven't had a problem that would disturb the convergence of this process.

the age of youngest child, the moments are given as:

$$\begin{aligned}
\frac{1}{J'_F}\sum_{j\in F}\left(\frac{s^{NCS}_{j'}(age<6)}{s^{NCS}_{j'}}-\frac{\int_{age<6}\left(1-\Pr\left\{\left\{d_{ij't}\right\}_{t=1}^{D_{j'}}=0\right\}\right)dV}{\int\left(1-\Pr\left\{\left\{d_{ij't}\right\}_{t=1}^{D_{j'}}=0\right\}\right)dV}\right) &= 0, \\
\frac{1}{J'_F}\sum_{j\in F}\left(\frac{s^{NCS}_{j'}(6\le age<17)}{s^{NCS}_{j'}}-\frac{\int_{6\le age<17}\left(1-\Pr\left\{\left\{d_{ij't}\right\}_{t=1}^{D_{j'}}=0\right\}\right)dV}{\int\left(1-\Pr\left\{\left\{d_{ij't}\right\}_{t=1}^{D_{j'}}=0\right\}\right)dV}\right) &= 0, \qquad (11)\\
\frac{1}{J'_F}\sum_{j\in F}\left(\frac{s^{NCS}_{j'}(age\ge 17)}{s^{NCS}_{j'}}-\frac{\int_{age\ge 17}\left(1-\Pr\left\{\left\{d_{ij't}\right\}_{t=1}^{D_{j'}}=0\right\}\right)dV}{\int\left(1-\Pr\left\{\left\{d_{ij't}\right\}_{t=1}^{D_{j'}}=0\right\}\right)dV}\right) &= 0,
\end{aligned}$$

and

$$\begin{aligned}
\frac{1}{J'_{NF}}\sum_{j\notin F}\left(\frac{s^{NCS}_{j'}(age<6)}{s^{NCS}_{j'}}-\frac{\int_{age<6}\left(1-\Pr\left\{\left\{d_{ij't}\right\}_{t=1}^{D_{j'}}=0\right\}\right)dV}{\int\left(1-\Pr\left\{\left\{d_{ij't}\right\}_{t=1}^{D_{j'}}=0\right\}\right)dV}\right) &= 0, \\
\frac{1}{J'_{NF}}\sum_{j\notin F}\left(\frac{s^{NCS}_{j'}(6\le age<17)}{s^{NCS}_{j'}}-\frac{\int_{6\le age<17}\left(1-\Pr\left\{\left\{d_{ij't}\right\}_{t=1}^{D_{j'}}=0\right\}\right)dV}{\int\left(1-\Pr\left\{\left\{d_{ij't}\right\}_{t=1}^{D_{j'}}=0\right\}\right)dV}\right) &= 0, \qquad (12)\\
\frac{1}{J'_{NF}}\sum_{j\notin F}\left(\frac{s^{NCS}_{j'}(age\ge 17)}{s^{NCS}_{j'}}-\frac{\int_{age\ge 17}\left(1-\Pr\left\{\left\{d_{ij't}\right\}_{t=1}^{D_{j'}}=0\right\}\right)dV}{\int\left(1-\Pr\left\{\left\{d_{ij't}\right\}_{t=1}^{D_{j'}}=0\right\}\right)dV}\right) &= 0,
\end{aligned}$$

where J'_F is the number of family movies in the NCS 2002 fall set, J'_{NF} is the number of the others, $s^{NCS}_{j'}(age<6)$ is the proportion of households with the youngest child less than 6 years old going to movie j' in any week, $s^{NCS}_{j'}(6\le age<17)$ is the proportion of households with the youngest child between 6 and 17 years old going to movie j' in any week, and $s^{NCS}_{j'}(age\ge 17)$ is the proportion of households with no child going to movie j' in any week.

6.2 Objective Function

We follow Berry, Levinsohn and Pakes (2004) and Petrin (2002) and generate the objective function and its optimal values and variances. We denote $G(\theta_0)$ by the population moment conditions that include the two sets of moments explained in 6.1. Because the population moment conditions can be assumed to uniquely equal zero at the true parameter values, θ_0, they can be written as:

$$E\left[G(\theta_0)\right]=0.$$

Then, the SMM estimator, θ_{SMM}, can be defined as:

$$\theta_{SMM}=\arg\min_{\theta\in\Theta}G^*(\theta)'G^*(\theta),$$

where $G^*(\theta) = A(\tilde{\theta})\hat{G}(\theta)$, $\hat{G}(\theta)$ is the sample analogue to $G(\theta)$, and $A(\tilde{\theta})$ is a consistent estimate satisfying that the inner-product of $A(\tilde{\theta})$ is equal to the inverse of the asymptotic variance-covariance matrix of the moments (obtained using $\tilde{\theta}$, a preliminary consistent estimate of θ_0).

The sources of variances in $V = E[G^*(\theta_0)G^*(\theta_0)']$ come from three independent sampling processes: the first arises from the process generating the product characteristics, the second from the consumer sampling process, and the third from the simulation process. Let $\Gamma = E[\partial G^*(\theta_0)/\partial\theta]$ denote the gradient of the moments with respect to the parameters evaluated at the true parameter values. Then, the asymptotic variance of $\sqrt{n}(\hat{\theta} - \theta_0)$ can be given as:

$$\left(\Gamma'\Gamma\right)^{-1}\Gamma'V\Gamma\left(\Gamma'\Gamma\right)^{-1}.$$

Finally, using the consistent estimates $\Gamma(\hat{\theta})$ and $V(\hat{\theta})$, the above asymptotic variance can be estimated.

7 Baseline Results and Fit

7.1 Goodness-of-Fit

To evaluate the out-of-sample fit by the model, we randomly leave out 36 movies from the dataset and estimate the parameter values with 200 movies.

Table 7 shows the estimates of both the consumers' demand and the studios' forecasting parameters. The first two panels of the table provide the estimates of the means and standard deviations of the households' taste parameters for observed movie characteristics. The third panel provides the estimates of the parameters to capture the additional effects of households with children that go to family movies. The fourth panel displays the estimates of the parameters for price sensitivities within three income groups. The fifth panel shows the estimates of the parameters of time-specific characteristics. The sixth panel provides the estimate of the parameter of the discount factor for advertising's effect. The last panel shows the estimates of the parameters to construct the studios' forecasts of the distribution of perceived qualities.

In the first panel, all estimates are positive, but the estimate for either the cumulative market shares or for winning an Academy Award is not significant. Cumulative market shares are circumstantial evidence of good quality, but they are not the dominant factor and the estimate is not significant. When a movie is nominated for an Academy Award, it receives a great deal of publicity. Thus, the additional gains from actually winning the award are not that significant. Also, most winners are no longer being shown in the theaters at the time the award is announced (they are closer to release in the home video market), so the estimate is not significant. Among the other estimates, the estimate for social learning is highly significant, reflecting the importance of the social learning process. In the second panel, all estimates are positive and significant. Relative to the t-statistics for the means of tastes, the t-statistics for the standard deviations of tastes for social learning, advertising, Oscar nomination and winning are all quite large. These estimates imply that there is heterogeneity among households' tastes for these four observed movie characteristics. They can explain why some people go to theaters at the beginning of a movie's release while the others do not. In the third panel, both estimates are positive and significant. In particular, the estimate for the family movie preference of the households in which the youngest child is older than 6 is not only higher but also more significant. The finding is consistent with the general belief that family movies are embraced by families with young children who can enjoy those films. In the fourth panel, the estimates are positive and significant. In the fifth panel, both estimates are positive, but the estimate for the holiday weekend dummy is significant, while that of the summer season dummy is not. In general, people are more likely to go to theaters during the summer, but they only watch a couple of blockbusters or some family movies. That may cause the estimate to be insignificant. In the sixth panel, the estimate is positive and significant.

The last panel shows how the studio forecasts the relationship between critics' ratings and the mean of the distribution for perceived qualities. The estimate for the constant is negative and significant: the estimate for the mean of the Metascores is positive and significant; the other two estimates are not significant. This finding implies that the top source for forecasting the word-of-mouth referrals can be critics' ratings.

7.2 Estimated Decay Patterns

Figure 2 shows the relationship of decay patterns from the data and those simulated from the model for two movies within the sample. The characteristics of a popular movie's decay pattern are a sharp decline at the beginning of its release and a smooth decay after the fourth week of its release. The estimated pattern for our example reflects these features well. The characteristics of an unpopular movie's decay pattern are that movie attendance does not decay sharply at the beginning of its release but that it moves toward \$0.00 more quickly. The estimated pattern of our example captures these aspects well.[15]

Figure 3 shows the relationship of decay patterns from the data and those simulated from the model for two movies out of sample. The patterns seem reasonably good, but the estimated patterns tend to better emphasize the features of the decay that stem from the movie's quality.

7.3 Model Fit

Table 8 shows the average decay patterns of high-rated movies and low-rated movies for first ten weeks after they are widely released, excluding the sleeper films. The second and fifth columns provide the average decay patterns found in the data. The third and sixth columns provide the average decay patterns estimated using the movies within the sample. The fourth and seventh columns provide the average decay patterns estimated using the movies out of the sample. In the case of low-rated movies, the estimated average pattern for the within-sample movies better approximates the pattern in the data, which implies that people are less inclined to go to theaters for low-rated films. That is, the low-rated movies are more influenced by the social learning factor. For out-of-sample movies, the average patterns for high- and low-rated movies are reasonably well fit to the patterns found in the data.

[15] Among widely released movies, there are some movies called "sleeper films," which exhibit a slow start but spark great interest and garner positive reviews down the line. For these movies, ticket sales follow a hump-shaped distribution over time, increasing at first and then decreasing. In the case of sleepers, they follow the platform release, (*i.e.,* they are released in a couple of major cities, such as New York or Los Angeles and then widely released throughout the U.S). We estimate the patterns for the period after such films are widely released.

8 Policy Implications

8.1 Comparison with HSX Prices

To see how well the model predicts the market demand, we compare the box office revenues for the first four weeks with those implied by HSX prices. HSX is an Internet-based simulated market in which online users buy and sell shares of their favorite movies and actors using a fake currency called H$. In valuing a movie's stock, users bid on the film's fourth-weekend cumulative box office revenue in the U.S. market, since its wide release. Remarkably, the stock price on the release date is an excellent proxy for the expectation about the movie, excluding word-of-mouth referrals from actual movie-viewers.

Figure 4 shows the relationship between the log of HSX prices on the release date and those of the first four weeks' cumulative box office gross as predicted by the model. The blue dots represent the movies with high Metascores; the black dots those with average Metascores; and red dots those with low Metascores. The black line in the figure shows the linear regression result. The estimate of the coefficient of the constant term is -0.2613 and its standard error is 0.0421. The estimate of the coefficient of the log of HSX prices is 1.0342 and its standard error is 0.0028. And R^2 is 0.6812. The model's forecasts are quite similar to the HSX prices on movies' release dates. But, HSX prices tend to underestimate for movies with low Metascores. Figure 5 shows the relationship between the cumulative box office gross for the first four weeks found in the dataset and those predicted by the model. The estimate of the coefficient of the constant term is -0.0864 and its standard error is 0.0069. The estimate of the coefficient of the log of HSX prices is 1.0149 and its standard error is 0.0005. Here, R^2 is 0.9241. Since the model predicts the social learning factors including word-of-mouth referrals and their influences on the ticket sales, the model's predictions for cumulative box office revenue are reasonably good and well fit to those found in the data.

8.2 Environment without Word-of-Mouth Referrals or a Learning Process

We consider a situation in which word-of-mouth referrals are not available, and uncertainty about a movie's quality is not resolved over time by a learning process. Here, consumers still use critics' ratings, but no longer acquire viewers' word-of-mouth referrals. Then, the prior distribution for a movie's quality is unchanged from the benchmark case, but the posterior distribution is different. Since the initial prior for a movie's quality cannot be updated by viewers' word-of-mouth recommendations, the posterior distribution is now identical to the prior distribution. This, in turn, affects the studio's optimization problem.

Figure 6 shows the relationship between the original advertising spending and the simulated optimal advertising spending in this hypothetical environment without learning. The dashed line is a 45-degree line (for reference) and the solid line is the regression line. The estimate of the coefficient of the constant term is -0.1573, and its standard error is 0.1217. The estimate of the coefficient of the original advertising spending is 0.9778, and its standard error is 0.0002. In this case, R^2 is 0.9583. To see how the changes in advertising spending differ by movie quality, we consider the following regression analysis:

$$\begin{aligned} ad_{j,sim} &= \underset{(0.1363)}{-0.3840} + \underset{(0.0002)}{0.9879} \cdot ad_j - \underset{(0.0150)}{0.2202} \cdot \Delta ir_j, \end{aligned}$$

where $ad_{j,sim}$ is the simulated optimal advertising spending for movie j in millions of dollars under the hypothetical situation, ad_j is the original advertising spending in millions of dollars and Δir_j is the difference between movie j's IMDb rating and the overall mean of IMDb ratings in our dataset. The numbers in parentheses are the standard errors and R^2 is 0.9590. This result shows that for good movies, producers should spend substantially more on advertising in a world with learning than they would in an environment without it. For bad movies, social learning makes much less difference for the optimal level of advertising expenditure.

Table 9 shows the optimal advertising spending and the total box office revenue in this environment as well as the benchmark case. In the world without learning, overall advertising volume diminishes by 3%, but the decline is greater for highly-rated movies. For good movies, studios spend 4.2% less on advertising without learning than they do with learning. For bad movies, learning makes a 1.4% decrease in the level of advertising expenditure. This result implies that the decreased multiplier effect is much influential on high quality movies, which is consistent with our preliminary finding that the multiplier effect is stronger for good movies. Interestingly, the box office revenue increases by 1.6% in general, even though the optimal volume of advertising decreases. For good movies, the box office gross increases by 2.5%, while there is a little change in box office performance for bad movies. Without learning, uncertainty about movie quality does not disappear, so that consumers receive a noisy signal about movie quality regardless of time. This makes some of moviegoers overestimating a movie's quality and induce them to watch the movie.

8.3 Environment with a More Accurate Prior and a Learning Process

We consider another hypothetical situation in which the distribution of critics' ratings is the same as that of viewers' perceived movie qualities. In this case, consumers get more accurate information about a movie's quality before its release. So, the initial prior distribution is now the true distribution of perceived qualities. However, critics' ratings are only small samples, so that uncertainty still exists at the time of the movie's release. Thus, the posterior distribution is updated by a learning process, as in the benchmark case. In this environment, the studio can more accurately predict future movie demand patterns, because it has better information at the moment that it chooses advertising spending. Such a scenario, again, changes the optimal level of advertising spending.

Figure 7 shows the relationship between the original advertising spending and the simulated optimal advertising spending under this hypothetical situation. The estimate of the coefficient of the constant term is 0.1523 and its standard error is 0.0693. The estimate of the coefficient of the original advertising spending is 0.9952 and its standard error is 0.0001. In this case, R^2 is 0.9766. To see how the effect differs by movie quality, we perform the following regression analysis:

$$\begin{aligned} ad_{j,sim} &= \underset{(0.0781)}{0.0133} + \underset{(0.0001)}{1.0014} \cdot ad_j - \underset{(0.0086)}{0.1350} \cdot \Delta ir_j, \end{aligned}$$

where R^2 is 0.9630. In this new environment, advertising spending increases for low-quality movies, while it decreases slightly for high-quality movies. Intuitively, when more accurate information about quality is available early on, the effectiveness of advertising for good movies increases, but its effectiveness for bad movies decreases. This induces less advertising for good movies and more for bad movies.

Table 10 shows the optimal advertising spending and the total box office revenue in this hypothetic situation as well as the benchmark case. In the world with more accurate initial prior, the overall advertising volume goes up by 0.2%. The advertising spending for good movies decreases by 0.2%, while that for bad movies increases by 0.8%. These changes are quite small, which is related to the correlation between critics and viewers' rating. Although critics do not perfectly reflect viewers' perceived qualities, they are good indicators of movie quality. Therefore, studios do not have to change the advertising strategy a lot. However, the change in box office revenue is relatively greater than that in advertising expenditure. In this case, consumers still have uncertainty in the very beginning, but the information itself is more accurate. Critics are inclined to praise quality of good movies more than moviegoers, whereas they tend to criticize quality of bad movies more harshly than consumers. These tendency cause that the box office revenue for a good movie decreases and that that for a bad movie increases.

8.4 Environment with Perfect Certainty

We consider the other situation with perfect certainty about movie quality. In this case, consumers know a movie's quality with perfect certainty before its release. In other words, consumers receive a single information about movie quality, which is not only identical but also most accurate. Then, the mean of the prior distribution is replaced with the mean of the true distribution of perceived qualities and the SD is simply 0. Also, the posterior distribution is the same as the prior distribution for all following weeks. In this environment, either word-of-mouth referrals or learning process does not have a role to reduce uncertainty, because consumers have the accurate information about movie quality regardless of time. Thus, the studio no longer takes advantage of consumers' ignorance of movie quality, which, in turn, influences the optimal volume of advertising spending.

Table 11 shows the optimal advertising spending and the total box office revenue in this situation as well as the benchmark case. In the world with perfect certainty, the overall advertising volume goes down by 2.7%, but the decrease is greater for good movies. The advertising spending for good movies declines by 3.6%, while that for bad movies diminishes by 1.4%. Similar to the findings in 8.2, this result is consistent with our preliminary finding that the multiplier effect is stronger for good movies. Generally, the total box office revenue decreases by 2.8%. The box office gross for good movies dwindles by 1.2%, while that for bad movies shrinks by 5.9%. With perfect certainty, consumers always receive the most accurate information about quality, which helps all of consumers estimating a movie's quality precisely. This effect is found greater for low quality movies.

9 Conclusion

This paper develops and estimates an equilibrium model of movie demand and the supply of advertising. The resulting estimates show how social learning can affect movie demand and studios' optimal advertising decisions. Also, social learning influences the process by which movie demand decays since the time of a film's release. The social learning channel can amplify or impede the effectiveness of advertising, depending on the movies' quality and the degree of uncertainty about quality. The empirical results show that, for the most part, social learning has a positive multiplier effect on advertising. The multiplier effect of social learning on advertising is found to be strongest for good movies. For this reason, movies with high critics' ratings and good word-of-mouth referrals receive the largest amount of advertising, whereas movies with low critics' ratings and bad word-of-mouth receive a smaller portion of the advertising budget. The simulation of the effects of social learning relative to a world without learning shows that for good movies, producers spend substantially more on advertising in an environment with learning than they would in an environment without it. For bad movies, social learning makes much less difference in determining the optimal level of advertising expenditure. The movie distributor's advertising expenditures are sensitive to both consumer uncertainty about movie quality and the speed with which potential movie-goers learn about movie quality.

References

Ackerberg, D. A. (2003). Advertising, Learning, and Consumer Choice in Experience Good Markets: An Empirical Examination, *International Economic Review* **44**(3): 1007–1040.

Beck, J. (2007). The Sales Effects of Word of Mouth: a Model for Creative Goods and Estimates for Novels *Journal of Cultural Economics* **31**(1): 5–23.

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

Berry, S., Levinsohn, J. & Pakes, A. (1995). Automobile Prices in Market Equilibrium, *Econometrica* **63**(4): 841–890.

Berry, S., Levinsohn, J. & Pakes, A. (2004). Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Vehicle Market, *Journal of Political Economy* **112**(1): 68–104.

Colussi, A. (2006). Migrants' Networks: An Estimable Model of Illegal Mexican Migration. Mimeo, University of Western Ontario.

De Vany, A. & Walls, W. D. (1996). Bose-Einstein dynamics and Adaptive Contracting in the Motion Picture Industry, *Economic Journal* **106**: 1493–1514.

Dellarocas, C. (2003). The Digitization of Word-of-Mouth: Promise and Challenges of Online Reputation Systems, *Management Science* **49**(10): 1407–1424.

Einav, L. (2006). Seasonality in the U.S. Motion Picture Industry, *RAND Journal of Economics* **61**(4): 631–653.

Elberse, A. (2007), The Power of Stars: Do Star Actors Drive the Success of Movies? *Journal of Marketing* **71**(4): 102–120.

Elberse, A. & Anand, B. N. (2007). The effectiveness of pre-release advertising for motion pictures: An empirical investigation using a simulated market, *Information Economics and Policy* **19**(3-4): 319–343.

Eliashberg, J., Elberse, A. & Leenders, M. (2006). The Motion Picture Industry: Critical Issues in Practice, Current Research, and New Research Directions, *Marketing Science* **25**(6): 638–661.

Erdem, T. & Keane, M. (1996). Deicision-making Under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets, *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.

Glaeser, E., Sacerdote, B. & Scheinkman, J. (1996). Crime and Social Interactions, *Quarterly Journal of Economics* **111**(2): 507–548.

Goolsbee, A., & P. J. Klenow (2002). Evidence on Learning and Network Externalities in the Diffusion of Home Computers, *Journal of Law and Economics* **45**(2): 317–344.

Gowrisankaran G. & M. Rysman (2007). Dynamics of Consumer Demand for New Durable Goods. Mimeo, University of Arizona.

Hall, G. & Rust, J. (2003), Simulated Minimum Distance Estimation of a Model of Optimal Commodity Price Speculation with Endogenously Sampled Prices. Mimeo, University of Maryland.

Ijiri, Y. & Simon, H. A. (1974). Interpretations of departures from the Pareto curve firm-size distributions, *Journal of Political Economy* **82**: 315–332.

Keane, M. & Wolpin, K. I. (1994). The Solution and Estimation of Discrete Choice Dynamic Programming Models by simulation and Interpolation: Monte Carlo Evidence, *Review of Economics and Statistics* **76**(4): 648–672.

Kopalle, P. K. & Lehmann, D. R. (1995). The Effects of Advertised and Observed Quality on Expectations About New Product Quality, *Journal of Marketing Research* **32**(3): 280–290.

Macieira, J. (2006). Extending the Frontier: A Structural Model of Investment and Technological Competition in the Supercomputer Industry. Mimeo, Virginia Tech.

Miller, D. & Shamsie, J. (2001). Learning Across the Life Cycle: Experimentation and Performance among the Hollywood Studio Heads, *Strategic Management Journal* **22**(8): 725–745.

Mobius, M. M., Niehaus, P. & Rosenblat, T. S. (2005). Social Learning and Consumer Demand. Mimeo, Wesleyan University.

Moretti, E. (2008). Social Learning and Peer Effects in Consumption: Evidence from Movie Sales, NBER Working Paper No. 13832.

Moul, C. C. (2005). *A Concise Handbook of Movie Industry Economics*, Cambridge, Cambridge University Press.

Moul, C. C. (2006). Consequences of Omitting Advertising in Demand Estimation: An Application to Theatrical Movies. Mimeo, Washington University.

Moul, C. C. (2007). Measuring Word of Mouth's Impact on Theatrical Movie Admissions, *Journal of Economics and Management Strategy* **16**(4): 859–892.

Nair, H. (2004). Intertemporal Price Discrimination with Forward-looking Consumers: Application to the US Market for Console Video-Games. Mimeo, Stanford University.

Petrin, A. (2002). Quantifying the Benefits of New Products: The Case of the Minivan, *Journal of Political Economy* **110**(4): 705–729.

Rennhoff, A. D. & Wilbur, K. C. (2007). The Effectiveness of Post-Release Movie Advertising. Mimeo, University of Southern California.

Santugini, M. (2006). Observational Learning in the Motion Picture Market. Mimeo, HEC Montreal.

Sood, S. & Dreze, X. (2006). Brand Extensions of Hedonic Goods: Movie sequel evaluations, *Journal of Consumer Research* **33**(3): 352–360.

Sorensen, A. (2007). Bestseller Lists and Product Variety, *Journal of Industrial Economics* **55** (4): 715–738.

Vogel, H. L. (2004). *Entertainment Industry Economics*, Cambridge, Cambridge University Press.

Table 1: Decision Factors to See a Movie

Reasons	Choose Yes (%)
The Cast	45
The Movie's Topic or Theme	39
TV (Net)	39
Trailer or Preview	20
Someone Recommended It	19
Good Reviews	18
Newspaper (Net)	16
Magazine Article	3

Data Source: *NAA* 2001 *MOVIE MARKET STUDY* available at *http://www.naa.org/display/movies/movie_study_print.html.*

Table 2: Summary of Data Sources

Types of Information	Sources
Aggregate Box Office Revenue	*Variety*
Total Advertising Expenditure	*Brandweek's Superbrand Report*
Critics' Ratings	*Metacritic*
Viewers' Ratings	*Internet Movie Database* (IMDb)
Production Budget Estimate	*Several Internet Movie Sites*
General Movie Information	*Variety*
Annual Average Ticket Price	*Box Office Mojo*
Census	IPUMS
Consumer Survey	NCS

Table 3: Critics' Ratings and Viewers' Ratings

	Critics' Ratings	Viewers' Ratings
Data	Metascores	IMDb User Ratings
Data Source	*Metacritic*	IMDb
Scale	0 to 100	0 to 10
Mean of Average	50.0	6.28
Mean of SDs	14.5	2.15

Note: Metascores and IMDb ratings use different scales, so we transform the Metascores to the same scale as the IMDb ratings.

Table 4: Summary Statistics of Movies

	Movies		
		Quality	
Mean	All	High	Low
Total Box Office Revenue ($Mln)	69.2	91.1	50.9
Opening-week's Box Office Revenue ($Mln)	26.9	35.6	22.8
Total Advertising Expenditure ($Mln)	22.9	25.5	20.8
Production Budget ($Mln)	47.1	54.8	42.9
Metascore	50.0	57.9	38.5
IMDb Rating	6.28	7.01	5.33
Duration (Wks)	15	16	13

Sources: The information is obtained from multiple sources in Table 2.

Table 5: IV Regression Results: Effects of Advertising Expenditure, Metascore and IMDb rating on the Log of Total Box Office Revenue

		Total Box Office Revenue							
							Quality		
							High	Medium	Low
Parameters	Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
a_0	*Constant*	2.106	1.645	2.291	1.588	2.295	3.574	3.262	0.477
		(0.034)	(0.049)	(0.843)	(0.055)	(0.879)	(0.911)	(2.353)	(0.514)
a_1	*Advertising*	0.758	0.728	0.431	0.726	0.382	0.784	0.733	0.460
		(0.004)	(0.005)	(0.119)	(0.002)	(0.129)	(0.010)	(0.013)	(0.028)
a_2	*Metascore*		0.084	-0.016	0.058	0.069	0.059	-0.070	0.078
			(0.002)	(0.020)	(0.008)	(0.059)	(0.012)	(0.010)	(0.005)
a_3	*IMDb Rating*				0.036	-0.083	-0.250	-0.092	0.320
					(0.003)	(0.056)	(0.016)	(0.058)	(0.035)
a_4	(*Advertising*)			0.045		-0.009			
	*(*Metascore*)			(0.003)		(0.008)			
a_5	(*Advertising*)					0.006			
	*(*IMDb Rating*)					(0.008)			
R^2		0.624	0.650	0.661	0.654	0.666	0.767	0.653	0.312

Note: The numbers in parentheses are the standard errors and the production budgets are used as instrument variables. # of observations: 236 movies in our data.

Table 6: OLS Regression Results: Effects of Social learning and Advertising on Weekly Decay Patterns

		Week			
Parameters	Variables	2	3	4	5
b_0	*Constant*	3.552	2.166	0.6025	-0.711
		(0.015)	(0.057)	(0.119)	(0.162)
b_1	*Advertising*	-0.0004	0.005	0.101	0.027
		(0.0001)	(0.0001)	(0.0001)	(0.0001)
b_2	*Metascore*	0.026	0.054	0.101	0.148
		(0.001)	(0.003)	(0.006)	(0.008)
b_3	*IMDb Rating*	0.026	0.121	0.176	0.184
		(0.001)	(0.003)	(0.006)	(0.008)
R^2		0.058	0.120	0.181	0.476

Note: The numbers in parentheses are the standard errors. # of observations: 236 movies in our data.

Table 7: Estimated Parameter Values

Parameters	Variables	Parameter Estimate	Standard Error
Means of Movie Characteristics	*Family Movie*	0.5785	0.2165
(β_k's)	*Action Movie*	0.3588	0.1761
	Cumulative Market Share	0.6548	1.2286
	Updated Movie Quality Info	0.7530	0.0041
	Adjusted Advertising Effects	0.1658	0.0751
	Oscar Nomination	0.3395	0.1752
	Oscar Winning	0.1227	0.1097
SDs of Movie Characteristics	*Family Movie*	0.5295	0.0198
(γ_k's)	*Action Movie*	0.3127	0.1167
	Cumulative Market Share	0.2302	0.1376
	Updated Movie Quality Info	0.1073	0.0020
	Adjusted Advertising Effects	1.9835	0.0171
	Oscar Nomination	2.9289	0.0102
	Oscar Winning	1.8114	0.1317
Family Movie Preferences	*Dummy for Youngest Child*		
(ζ_k's)	*Under 6*	0.0334	0.0229
	Between 6 and 17	0.0563	0.0130
Price Sensitivity of			
1^{st} Income Cohort, α_1	*Price to Watch a Movie*	8.3002	0.1249
2^{nd} Income Cohort, α_2	*Price to Watch a Movie*	5.9757	0.0146
3^{rd} Income Cohort, α_3	*Price to Watch a Movie*	3.8889	0.0014
Time-Specific Characteristics	*Holiday Weekend Dummy*	0.1514	0.0550
(ρ_k's)	*Summer Season Dummy*	0.0634	0.1033
Discount Rate of Ad Effect			
(η)		0.3000	0.0023
μ_l's	*Constant*	-0.9964	0.3671
	Standardized Media Critic	1.5963	0.7218
	(*Standardized Media Critic*)^2	-0.0686	0.0753
	Production Cost	0.0087	0.0096

Note: For estimation, both the movie-specific fixed effects and the time-specific unobserved heterogeneities are included.

Table 8: Model Fit

Week	High-rated Movies			Low-rated Movies		
	Data	Within Sample Estimate	Out of Sample Estimate	Data	Within Sample Estimate	Out of Sample Estimate
1	0.0224	0.0250	0.0261	0.0129	0.0140	0.0174
2	0.0117	0.0085	0.0096	0.0066	0.0054	0.0064
3	0.0069	0.0058	0.0055	0.0038	0.0038	0.0041
4	0.0043	0.0034	0.0036	0.0021	0.0022	0.0030
5	0.0027	0.0019	0.0020	0.0011	0.0011	0.0015
6	0.0018	0.0010	0.0013	0.0006	0.0006	0.0010
7	0.0011	0.0006	0.0009	0.0003	0.0003	0.0006
8	0.0007	0.0005	0.0006	0.0002	0.0002	0.0004
9	0.0005	0.0004	0.0005	0.0001	0.0002	0.0003
10	0.0003	0.0003	0.0004	0.0001	0.0001	0.0003

Table 9: Comparison between Benchmark and a World without either WOM Referrals or Learning

	Movies		
		Quality	
Mean	All	High	Low
Advertising Spending *in a new Equilibrium* ($Mln)	21.63	23.21	19.82
Advertising Spending *in benchmark* ($Mln)	22.31	24.23	20.10
Difference ($Mln)	-0.68	-1.02	-0.28
Difference (%)	-3.05	-4.21	-1.39
Box Office Revenue *in a new Equilibrium* ($Mln)	65.35	82.20	45.98
Box Office Revenue *in benchmark* ($Mln)	64.29	80.21	45.97
Difference ($Mln)	1.06	1.99	0.01
Difference (%)	1.65	2.48	0.02

Table 10: Comparison between Benchmark and a World with both More Accurate Initial Prior and Learning

	Movies		
		Quality	
Mean	All	High	Low
Advertising Spending *in a new Equilibrium* ($Mln)	22.35	24.18	20.26
Advertising Spending *in benchmark* ($Mln)	22.31	24.23	20.10
Difference ($Mln)	0.04	-0.05	0.16
Difference (%)	0.17	-0.21	0.80
Box Office Revenue *in a new Equilibrium* ($Mln)	63.44	78.26	46.39
Box Office Revenue *in benchmark* ($Mln)	64.29	80.21	45.97
Difference ($Mln)	-0.85	-1.95	0.42
Difference (%)	-1.32	-2.43	0.91

Table 11: Perfect Certainty Vs. Uncertainty

	Movies		
		Quality	
Mean	All	High	Low
Advertising Spending *with Perfect Certainty* ($Mln)	21.70	23.35	19.81
Advertising Spending *in benchmark* ($Mln)	22.31	24.23	20.10
Difference ($Mln)	-0.61	-0.88	-0.29
Difference (%)	-2.73	-3.63	-1.44
Box Office Revenue *with Perfect Certainty* ($Mln)	62.52	79.26	43.25
Box Office Revenue *in benchmark* ($Mln)	64.29	80.21	45.97
Difference ($Mln)	-1.77	-0.95	-2.72
Difference (%)	-2.75	-1.18	-5.91

Figure 1: Weekly Box Office Performances

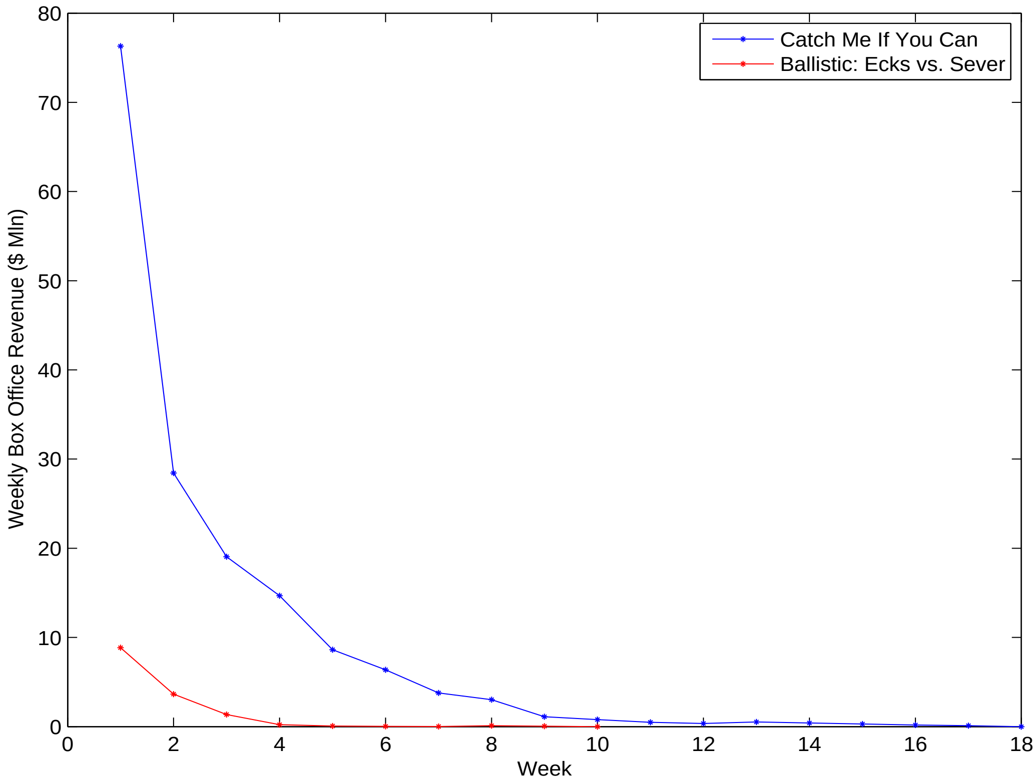

Figure 2: Decay Patterns: Within Sample Fit

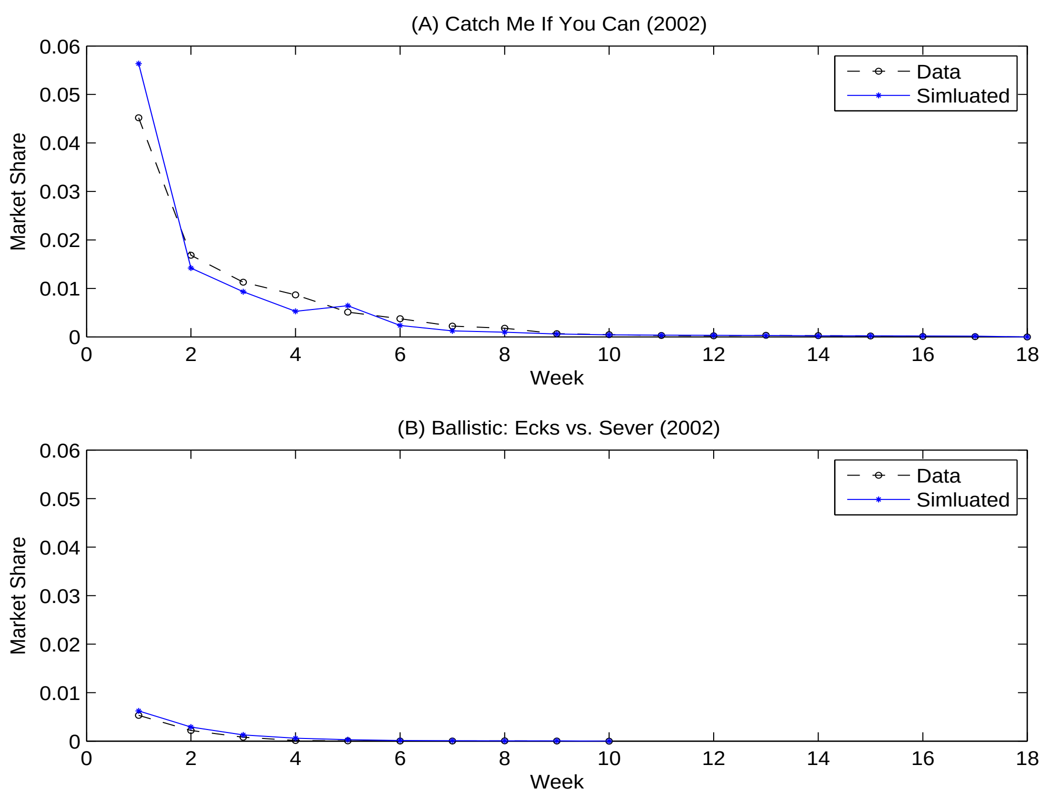

Figure 3: Decay Patterns: Out of Sample Fit

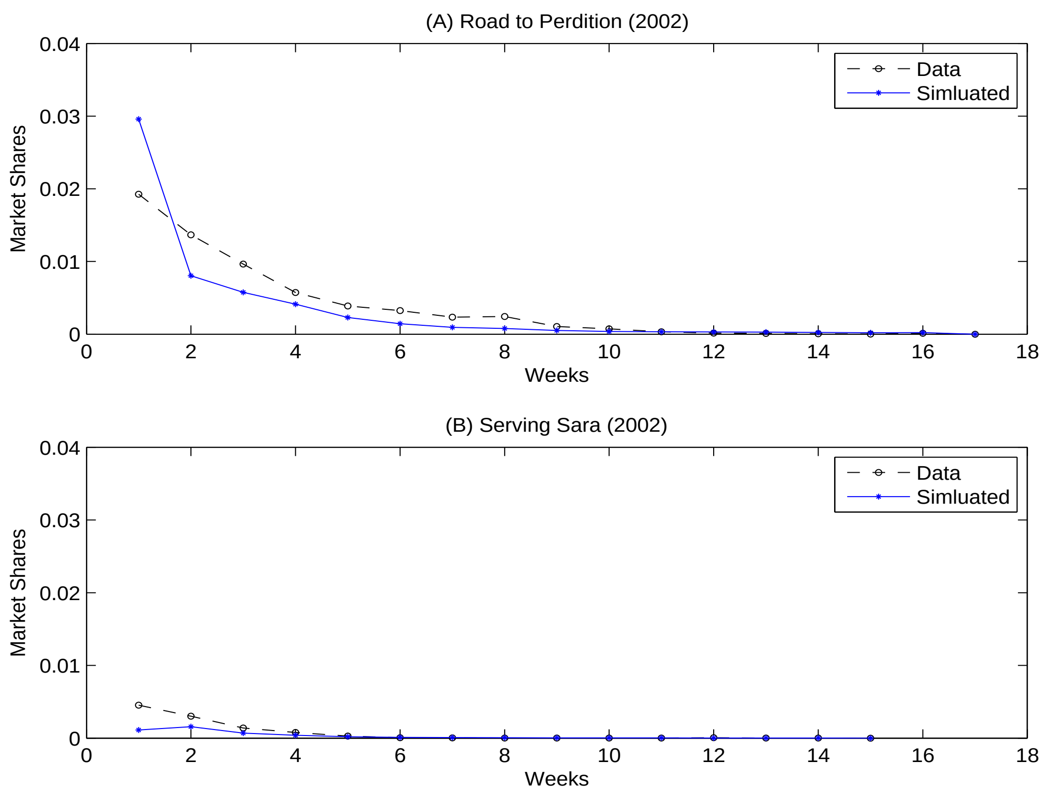

Figure 4: Comparison Between the Model's Forecasts and HSX Prices

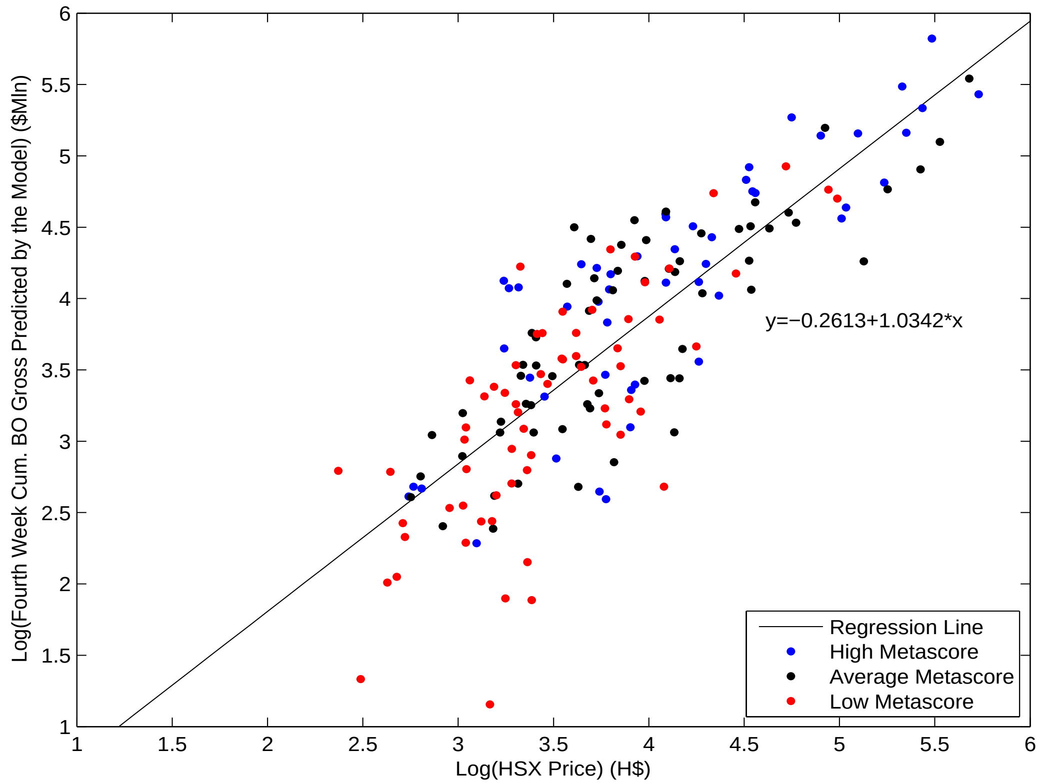

Figure 5: Comparison Between the Model's Forecasts and the Fourth Week Cumulative Box Office Revenues

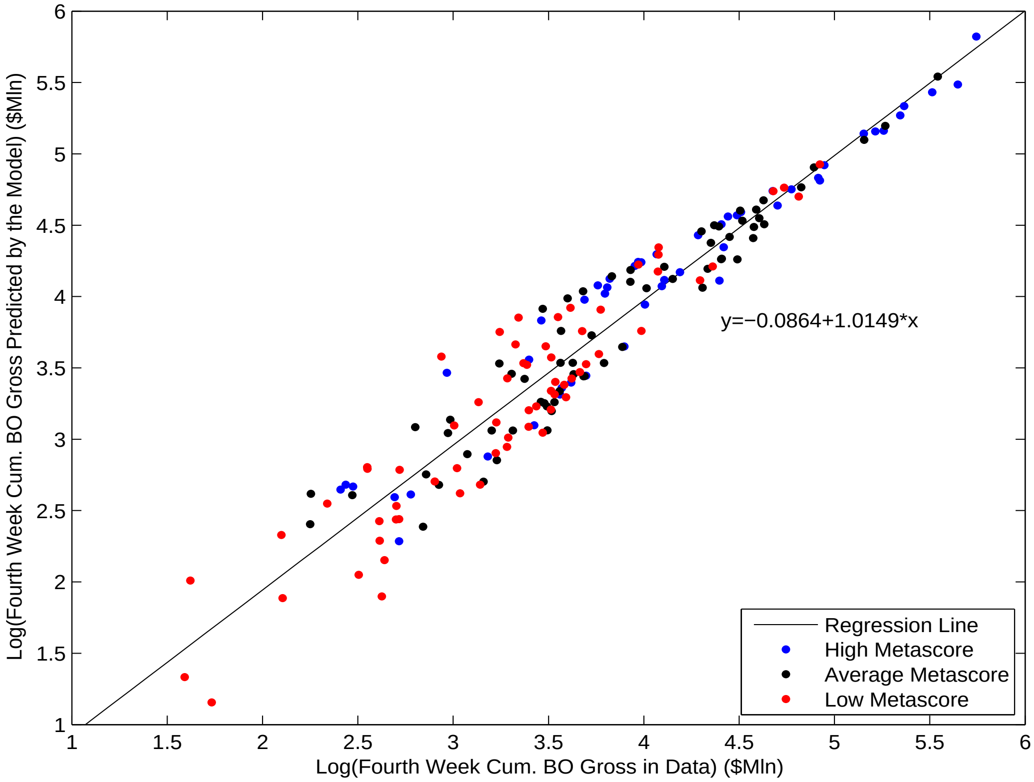

Figure 6: Comparison Between the Original Advertising Spending and the Simulated Advertising Spending: Under an Environment without Word-of-Mouth Referrals

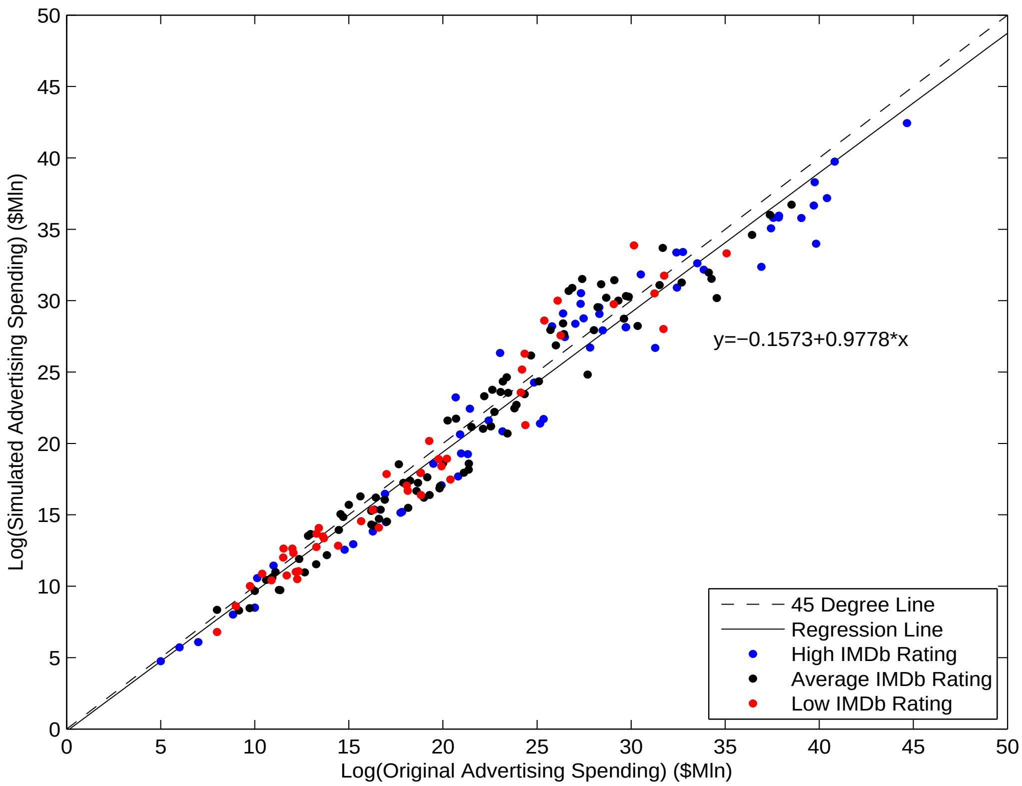

Figure 7: Comparison Between the Original Advertising Spending and the Simulated Advertising Spending: Under an Environment with a Faster Learning Process