A Structural Model of Advertising Signaling and Social Learning: The Case of the Motion Picture Industry

Haiyan Liu†
University of South Florida
April 23, 2015

Abstract

This paper analyzes how social learning among consumers shapes the optimal strategies of firms in the motion picture industry for signaling product quality through advertising. I analyze the distribution of advertising spending over time with a structural equilibrium model that incorporates both pre-release information asymmetry and post-release consumer learning. Two roles of informative advertising, reaching and signaling roles, are put into one framework to be empirically separated and quantified. I estimate the model using weekly data on advertising spending and movie characteristics from movie theater admissions in the United States. By estimating studios’ equilibrium advertising policy function, I demonstrate that advertising does play a quality signaling role in the movie industry. I also evaluate the pre-release information uncertainty for firms and consumers, respectively, and how the information asymmetry is reduced separately by the signaling effect of advertising and by Word-of-Mouth. Counterfactual experiments are performed to distinguish the amount of money used for purposes of signaling product quality as well as reaching consumers. I also discuss studios’ advertising strategies under different market information structures.

Keywords: Information Asymmetry, Consumer Learning, Signaling Advertising, Motion Picture Industry

JEL: D22, D82, D83, L15, L82, M37

*This paper is based on my dissertation. I am grateful to my committee members, Steven Stern, Simon Anderson, Natasha Zhang Foutz and Federico Ciliberto for their generous guidance and support. This work has benefited from conversations with Nathan Larson, David Mills, Michelle Sovinsky Goeree, Ginger Jin, John Rust, Christopher T. Conlon, and Yiyi Zhou, as well as comments in the IO research meetings at UVA and participants. I acknowledge the Bankard Fund for Political Economy for financial support. I am solely responsible for any errors.

†University of South Florida, Department of Economics, hliu4@usf.edu
1 Introduction

Many markets are characterized by the information asymmetry between firms and consumers. For new products especially, consumers are motivated to learn about product quality from all possible credible information sources in order to differentiate high-quality products from low-quality products. On the other hand, firms with high-quality products are also motivated to send “signals” about their product quality to influence consumer learning. Advertising can be one of those quality signals that may avoid the lemons problem, i.e. the problem of low-quality firms outpacing high-quality firms.

The so-called “money-burning” theory\(^1\) of advertising was introduced by Nelson (1974), formalized by Kihlstrom and Riordan (1984), and Milgrom and Roberts (1986). This theory works according to the following mechanism. Suppose that consumers cannot directly ascertain product quality and, therefore, run a risk of buying inferior products. The conspicuously-expensive advertising (e.g. in the Super Bowl) of high-quality products may distinguish those products from the others because a high-quality product would garner repeat purchases, which would allow the firm to recoup its advertising spending. Firms with low-quality products would not recover such an investment because consumers, if fooled into buying the first time, would not buy again. The interaction of experience and repeat purchases can create an asymmetry in the returns to advertising and, therefore, support the signaling equilibrium.

Although it is well established in theory that advertising can be used as a signal of product quality, little empirical analysis has tested or measured this signaling effect, especially in an equilibrium setting between consumers and firms. In this paper, I empirically study how word-of-mouth (WOM)\(^2\) communications among consumers supports the signaling effect of advertising in the context of consumer learning through others’ consumption experience. This type of learning is more important for products that are purchased infrequently, such as entertainment goods and durable goods. For industries in which repeat purchase by the same consumer is unlikely, WOM actually substitute for repeat purchases to support the signaling effect of advertising when communication cost is sufficiently low. It is worthy noted that the Nelson-Milgrom-Roberts analysis is better suited to explain the marketing strategies of non-durable products, because it mainly focuses on advertising signaling in the context of consumer learning through personal consumption experience in markets where repeat buying happens frequently.

To study how WOM communications among consumers shapes firms’ optimal strate-

---

\(^1\)In theory, “money burning” advertising means firms “burn” money just to show they can afford it and the advertising need not to have direct informative content. In my paper, it means studios spend extra more money to show they are very confident about their movies’ unobservable quality than just to inform consumers about the existence and observable attributes of their movie.

\(^2\)Word-of-mouth (WOM) communication is one type of social learning among consumers. “Social learning” includes other types of learning, such as observational learning. In this paper, I mainly focus on WOM and use two terms interchangeably.
gies of using advertising as a signal of product quality, I propose a structural equilibrium model to describe both consumers’ and firms’ decisions under uncertainty about product quality. In this paper, the data used for estimation comes from widely released movies from 2000 through 2005 in the U.S. theatrical market. This particular industry provides an ideal test-bed for the following reasons. First, there are enough observations of similar circumstances to enable a broad enough dataset for the empirical work. New products (movies) of uncertain quality (entertainment value) to consumers are introduced to the market every week. Second, studios routinely use marketing research to gauge the overall quality perceptions of their new movies prior to release (Turner and Emshwiller 1993, Wall Street Journal), thereby producing an information asymmetry between studios and consumers. Third, in most industries, the ability to signal product quality can come through several channels, such as low introductory prices and product warranties. Since prices of movie tickets are the same regardless of movie quality, price signaling is ruled out. Fourth, both advertising and WOM play very important roles for information learning in this industry. Pre-release quality uncertainty and post-release social learning are the two main factors for studios to consider when making their advertising decisions.

In this paper, I contribute to the informative advertising literature by empirically distinguishing between the reaching effect (with direct information) and the signaling effect (with indirect information) of advertising. Before consumers decide whether or not to purchase a new introduced product, such as a newly released movie, they need to collect two types of information. First, consumers need to be aware of a new product coming as well as its observable attributes through advertisements; hence advertising plays its reaching role. However, more importantly, consumers are motivated to learn about the product’s unobserved quality. If consumers infer the product quality from the observed advertising strategies of the firm, then the advertising plays its signaling role. On the other hand, when the firm of the new product decides its optimal advertising spending, the firm also needs to consider two roles of advertising: how many consumers the advertising can reach and to what extent its confidence on the product can be shown through advertising. The reaching and signaling roles of advertising has been theoretically discussed in the literature that is related to informative advertising, however, to the best of my knowledge, my study is the first one that put those two roles of advertising into one framework to empirically separate and quantify them.

The intuition of separately identifying those two roles of informative advertising is described as following. For consumers who enter the market at different time, they should have different information sources. For example, after a new movie is released, consumers who enter the theaters during the opening week are primarily influenced by the firm-generated information (advertising). However, consumers who enter the market in the post-release weeks are influenced largely by social learning (WOM). Therefore, advertising in the pre-release stage has both a signaling effect and a reaching effect on
demand, while advertising in the post-release stage only has a reaching effect. The changes in the information structure and advertising spending over time help distinguish between those two informative effects of advertising.

The signaling theory is usually described as a situation where there are just two types of firms, advertising spending is observed by all consumers, and repeat purchases drive the motive to signal quality. In any actual empirical market, these conditions are unlikely to be met in a pure form. In this paper, I have built up a structural model which draws heavily from the theoretical literature of informative advertising but attuned to the dataset under consideration. I propose an equilibrium model which analyzes both consumer learning process and studios’ optimal allocation of advertising spending over time. The features of my model are summarized in the following five aspects.

First, instead of considering only low-or-high-quality movies, I consider many possible quality types, with higher quality corresponding loosely to a larger number of people who will consume. This leads to a Bayesian-Nash equilibrium in which firms advertise more heavily on higher quality movies in the separating equilibrium.

Second, the ways that advertising can affect the demand of a movie are modeled in detail. Following Butters (1977), I assume that advertisements are sent out as a series of messages after studios decide their advertising spending budgets. When a consumer receives at least one advertisement, she is reached by the studio and aware of this movie. The more advertisements this consumer receives, the better the movie is inferred by her in the separating Nash equilibrium. Therefore, advertising can be used both to reach consumers and to signal product quality.

Third, I divide the time into two periods: pre-release weeks (including the opening week) and post-release weeks. This simplified two-period model helps reduce the computation burden, but still capture the information structure of this market.

Fourth, the information learning processes of both studios and consumers are modeled. The studio receives a noisy signal about its movie’s true quality and updates its belief in the Bayesian learning framework. Before the release of a movie, consumers receive advertisements and update their beliefs about a movie’s quality. After the movie is released, WOM becomes a credible but noisy signal of quality, and potential new consumers update their beliefs again. The more and faster that WOM communication occurs, the more accurate the quality signal revealed to consumers is.

Fifth, on the supply side, the studio chooses optimal advertising spending for each period to maximize expected profit which is written in the Bellman equation format. On the demand side, consumers make static discrete choice about whether or not to watch this movie conditional on being reached by advertising. Therefore, the probability that a consumer decides to watch a movie is composed of two parts: the probability that she is reached by the advertising and the probability she is convinced to watch the movie.

Because it is infeasible to access complete and reliable data on studios and consumers’
private information, I use this game-theoretic model to recover the unobserved information that is consistent with the observable data on consumers’ choices and studios’ actions. Instead of estimating the demand and the supply parameters separately, I estimate all structural parameters jointly by using detailed data including movie characteristics, market performance, and studios’ weekly advertising spending in the movie theater market from 2000 through 2005. Since studios’ optimal pre-release advertising policy function is an equilibrium result of the incomplete information game between studios and consumers, it cannot be written in an analytical format explicitly. Therefore, I use the Chebyshev approximation to approximate it. In addition, instead of maximizing the likelihood function directly, I take equilibrium outcomes of the model as constraints and use the MPEC (Mathematical Programming with Equilibrium Constraints) method of Su and Judd (2012) to simplify the estimation. The significant advantage of the MPEC method over other methods, such as full information maximum likelihood (FIML) method, is that it does not require computations of the equilibrium to the model repeatedly during estimation.

The estimated advertising policy function, as an increasing function of unobserved quality conditional on observed characteristics of a movie, supports the existence of the signaling equilibrium. I first estimate the specification with consumers who do infer quality information from advertising and then compare with the specification with consumers who do not infer quality information from advertising. Comparing the maximized likelihood values, the first specification is preferred, which demonstrates the existence of the signaling effect of advertising in this industry. The estimated information parameters (prior variances and posterior variances of expected movie quality) from my model also show that studios usually do not learn about movie’s true quality very precisely, and WOM is a much more efficient channel for consumers to learn the true quality of a movie. In the post-release weeks, the uncertainty about a movie’s quality is reduced by more than 90% mainly through the WOM channel.

After estimating the structural parameters, I conduct a set of counterfactual experiments to separately quantify the singling and reaching effects of advertising. In the simulated cases, advertising is only used to reach consumers, without any signaling effect, and the optimal advertising spending problems are solved for the studios in my sample. The simulated total advertising spending for all movies in my sample is around $9.5 billion which is only 73% of the case when advertising is used for both signaling and reaching. This means that around 27% of advertising spending for movies in my sample is “burned” for the signaling purpose, while 73% of the advertising money is spent to reach consumers.

Using the same simulated results, I study studios’ optimal strategies on allocating advertising spending over time. In the case when advertising is only used to reach consumers, on average, advertising money is arranged much more evenly over time, with around 50%
spent in the pre-release stage and another 50% spent in the post-release stage. However, in the case where advertising plays both signaling and reaching roles, studios actually allocate around 76% of advertising money in the pre-release stage, in order to achieve the signaling purpose. The counterfactual experiments also show that information revealed by both advertising signaling and WOM even prevents movies with very low quality from entering the market. When word-of-mouth communication has lower cost and become more efficient, less advertising spending is required for the signaling purposes.

The rest of the paper is organized as follows: In section, I review the literature most closely related to my work. In section 3, I briefly describe the U.S. movie theater market and the data used in this empirical exercise. Section 4 lays out the model, defines the pure strategy Nash signaling equilibrium. Section 5 explains my empirical strategy, followed by a brief discussion of identification. Section 6 presents the estimation results. Section 7 conducts the counterfactual experiments. Section 8 concludes the paper with some discussion about future work.

2 Related Literature

My research draws heavily from the theoretical literature of informative advertising. Since Nelson (1970) first made the important distinction between search goods and experience goods, the literature that is related to informative advertising can be divided into two groups. One focuses on how advertising conveys “hard” (direct) information about a product’s existence and attributes. Butters (1977) offers the first equilibrium analysis of informative advertising. In his model, advertising is used to convey information on product existence and price in the context of monopolistic competition. All firms offer the same product without horizontal or vertical differentiation, but have informational differentiation\(^3\) after advertising. The framework used in this paper to study how advertising reaches consumers is grounded in Butters-type model. Later, Grossman and Shapiro (1984) use the similar framework to analyze how advertising reaches consumers, but allow product differentiation along two dimensions (both information and location). Anderson and Renault (2006) extend the directly informative advertising literature by discussing firms’ choices of the type of information transmitted in advertisements.

The other group of papers focuses on how advertising conveys “soft” (indirect) information, from which consumers can correctly infer unobservable quality of products. Nelson (1974) argues that the interaction of experience and repeat purchases can create an asymmetry in the returns to advertising and, therefore, supports the signaling equilibrium. Kihlstrom and Riordan (1984) and, later, Milgrom and Roberts (1986) formalize

\(^3\)Informational differentiation means products produced by two firms that are identical in other respects may still be differentiated in the eyes of a consumer because her information set about the product differs from that of the other.
Nelson’s intuition with consumer rationality. Hertzendorf (1993) extends the analysis by allowing consumers to observe the monopolist’s advertising expenditure with error.

My paper is inspired by those theories; however, it presents a model from empirical perspectives. More importantly, in my model, advertising serves as a mechanism by which awareness is raised and product quality is signaled. Therefore, advertising spending is not simply a dissipative expense. By separating the reaching role and signaling role of advertising, I can tell how much more money studios need to spend in order to get the signaling equilibrium. That extra money studios spend is called the “burned” money for signaling purpose in this paper. Following Hertzendorf (1993), I also consider the possibility that consumers observe the monopolist’s advertising expenditure with error instead of observing the advertising spending directly. Compared to most theoretical papers that discuss the interaction of price and advertising in the signaling framework, I can only focus on advertising because of the movie industry’s uniform pricing feature.

Early empirical work about informative advertising mainly examines the relationship between price, quality and advertising. Studies that look for evidence of the signaling effect of advertising mainly focuses on the detection of a positive relationship between advertising spending and product quality, so as to indirectly support the theory of signaling advertising. There are several main reasons that direct tests of advertising’s signaling effect are difficult. First, it is hard to tell whether advertising conveys hard information, soft information, or both. Second, other possible signals of quality, such as low introductory price, may interact with advertising, which complicates the analysis. Given the focus on correlation with quality, these studies must quantify quality measures, which is extremely problematic. They also suffer from industry heterogeneity when using cross-industry data.

Thomas, Shane, and Weiglet (1998) improve the literature by investigating data from the U. S. automobile industry. They find that car models that are priced higher than the full information price level tend to have greater advertising levels. Such positive relationships are weaker for older car models, about which consumers are already well informed. Therefore, they conclude that manufacturers use both price and advertising to signal the quality of their products. Horstmann and MacDonald (2003) provide a related analysis that focuses on the compact disc player market. By employing panel data, they avoid constructing a quality index, instead, they examine whether the time-series behavior of price and advertising is consistent with the prediction of signaling advertising models. They find that the observed firm advertising and pricing behavior is inconsistent with the predictions of signaling model of advertising.

Erdem and Keane (1996) develop a structural model of household information learning behavior in the laundry detergent market. In their model, consumers have imperfect

---

4Dissipative advertising does not directly affect demand (neither persuasive nor directly informative content) and it is easy to observe that a substantial amount of money has been spent.
information about a brand’s attributes and learn through both consumption experience and advertising exposure. This paper provides a framework to analyze how the content of each advertising message provides noisy but “direct”\(^5\) information about brand attributes. Based on their 1996’s paper, Erdem, Keane and Sun (2008) add price and advertising frequency/intensity as two more information sources for consumers. Both price and advertising intensity are assumed to be linear functions of brand quality level, and consumers infer the quality level of a brand from these two “indirect” signals. They find that advertising frequency does influence consumer learning, although it is less quantitatively important than price. Ackerberg (2001) and Ackerberg (2003) empirically studies consumer learning behavior in the retail yogurt market by allowing that advertising intensity has both “informative” effect and “persuasive” effect\(^6\) on demand. He finds that, in this market, advertising is primarily used to inform consumers instead of persua ding consumers.

Those above-mentioned four papers use very similar models to investigate how advertising influence demand, but from different perspectives. The first two papers focus on the informative effect of advertising on the demand, while the last two papers focuses on distinguishing the informative and persuasive effects of advertising. Compared to those above-mentioned papers, my study considers both demand side and supply side within one framework. I focus on how the interaction of advertising and WOM support the signaling role of advertising in an equilibrium setting. Two kinds of informative roles of advertising are separately modeled and empirically quantified. Both Erdem, Keane and Sun (2008) and Ackerberg (2003) assume that advertisement intensity is a linear function of unobservable attributes and analyze how consumers infer attributes from observed advertisement intensity. While, in this paper, advertisement intensity as a function of observable attributes and unobservable quality is solved as an equilibrium results.

In previous literature, consumers are assumed to know the existence and observable characteristics of a brand. Goeree (2008) relaxes this assumption by estimating a model of limited consumer information, where advertising influences the set of products from which consumers choose to purchase. She investigates the PC industry and finds estimated markup is much higher than that predicted by full information models. She also finds estimated demand curve is less elastic than traditional models, because the market is less competitive when consumers are not fully informed. However, in her paper, there is no uncertainty about product attributes from consumers’ perspective. In my model, advertising intensity is used both to inform consumer the existence and observable attributes of a movie (reaching effect), and to signal unobservable movie quality (signaling effect).

\(^5\)Here, “direct” signal means consumers learn directly from signals instead of infer from signals.

\(^6\)Persuasive effect means advertising primarily affects demand by changing consumer taste and creating brand loyalty. In those empirical studies, advertising spending is considered as part of the utility function.
There is a growing literature on advertising and social learning using data from the movie industry in both economics and marketing areas. Basuroy, Desai, and Talukdar (2006) discuss the interaction of sequels and advertising expenditures as signals of movie quality using a reduced-form analysis method. Elberse and Anand (2007) use data from the Hollywood Stock Exchange to study the impact of movie advertising on expected revenue. They find that advertising has a positive and statistically significant impact on market-wide expectations prior to release; this impact is less for lower-quality movies. My work explains the mechanism supporting this positive relationship. Santugini (2007) sets up a dynamic structural demand model to investigate how consumers learn about the quality of a movie by observing its market share during its release week. I extend his work by allowing information about movie quality to be revealed by both market share (driven by advertising) and WOM. Moul (2007) and Moretti (2011) identify the impact of social learning on a movie’s sales, without considering how social learning affects the supply side’s strategic choice of advertising. Joo (2009) considers studio advertising and consumer learning in the equilibrium setting. However, in her model, advertising affects the consumer utility function directly, and studios are assumed to do all advertising before the movie’s release. In my model, advertising affects consumer information sets, and the observed large share of advertising expenditures in the pre-release stage is an endogenous choice.

3 Industry Background and Data

The theatrical motion picture industry has an economic importance in the global economy and U.S. economy. In 2013, global box office for all films released in each country around the world reached $35.9 billion, and U.S. (and Canada) box office was around $10.9 billion. More than two-thirds of the U.S./Canada population (68%) – or 227.8 million people – went to the movies at least once in 2013. The regular moviegoer segment which is defined as the segment of U.S. population who see at least 6 movies a year in cinemas currently is 35% of the U.S. population (MPAA 2013). A movie can recoup its investments from both theatrical windows (both local and global theatrical markets) and nontheatrical windows (such as home video market, pay television, network television, video game and merchandising). Among those numerous revenue windows, the theatrical box-office revenue is believed to be the most important performance metric for distributors, since it is also an indicator of the movie’s potential sales in other distribution windows.

Hollywood is a big spender on advertising. According to Nielson Monitor-Plus, movie studios spent $3.734 billion to buy advertising in the United States; movies ranked fifth in the nation among paid advertising categories in 2007. Advertising spending also constitutes a large share of a movie’s budget. For example, in 2007 the average production budget for a theater-released movie from a major Hollywood studio reached $70.8 mil-
lion, and studios spent another $35.9 million on marketing that movie to the public (www.mpaa.org).

New movies enter the theater market every week and exit the market after a few weeks. Due to the short life cycle of new movies and their uniqueness as typical experience products, the motion picture industry is characterized by information uncertainty problems. Both supply and demand sides are involved in active information learning to reduce their uncertainties. One key risk studios need to cope with is called “performance risk” that is how the market perceives and reacts to a new movie after its release. In recent years, sequel movies become especially prevalent, which may reflect studios’ eagerness to emphasize on the well-established properties of movies to better manage the performance risk. For the demand side, each movie is unique and the quality of a new movie is also ex ante uncertain. Consumers do not know for sure whether they will like the movie or not before they actually go to the theater and watch it. Therefore they make their watching decision based on observable characteristics of the movie such as the director, actors, the genre and ratings, and they also learn about the unobservable quality from different information sources such as “firm-generated” information from movie studios and “consumer-generated” information from their peers.

Although it is very difficult to accurately predict revenues and profits of new movies, studios/distributors\(^7\) often conduct formal market research for movies which are expected to hit more than six hundred theaters. In general, studios should have more information than consumers, as a result. Thorough market screenings and surveys are commonly used by studios; hence, it is hard to imagine that studios simply follow the “50%” rule of thumb\(^8\) with their advertising budget. In my data sample, total advertising spending averages around 50% of the production budget, but the ratio between those two varies widely across movies. It ranges from 1.5% to 87.5%, with a standard deviation of 65.5%. Figure 1 further shows that although the per-release advertising spending is highly correlated to the production budget, the post-release advertising has much lower correlation to that budget. Therefore, studios are making prudent decisions on advertising spending, and they respond to the market fairly quickly when critics and moviegoers disseminate feedback about movie quality. However, mistakes are to be expected. In the data, we observe that the size of an advertising spending does not always directly correlated with box office revenue. For example, the total advertising spending for the movie “I Spy” released in 2002 was more than $45 million, yet it only generated less than $34 million box office revenue during its 12 weeks in theaters.

Another interesting phenomenon existing in the U.S. theatrical movie market is the

---

\(^7\)In this paper, I use “studios” and “distributors” interchangeably. In practice, major studios do play the role of distributor, while independent distributors tend to fill market segments which are not covered by majors.

\(^8\)“50%” rule of thumb means a movie’s advertising budget should be 50% of its production budget. If a movie costs $100 million to make, the studio needs an additional $50 million to sell it.
contrast between the allocation of box office revenue and advertising spending over time. In my sample, the average box office revenue from a movie’s opening week is $14.13 million, while the average box office revenue in weeks following the opening week is $33.40 million. On average, around 25% of a movie’s box office revenue comes from the opening week, yet about 75% of a movie’s total advertising budget is spent in the pre-release stage. About $5 million, on average, was spent in the weeks after the release, compared to $15.7 million, on average, in the weeks before a movie’s release and its opening week. This contrast between box office revenue and advertising spending raises questions about advertising’s effects on demand over time and how studios dynamically decide their advertising spending.

After the opening week, advertising decreases quickly over time, and WOM (social learning) among consumers becomes the main quality information source. With the emergence of social media such as Twitter and Facebook, social learning plays an increasingly important role in the movie industry today. The impact of social learning is reflected by the fact that sales trends for movies diverge over time after their release. One important question is how quickly social learning reveals movie quality to a potential consumer. Figure 2 shows the sale trends of “Bruno” and “District 9,” which were released in the summer of 2009. These two movies’ decay patterns diverged after their releases, but especially did so from week one to week two. The weekly box office revenue of “District 9” dropped about 5 percent, from $37 million to $35 million, while the weekly box office revenue of “Bruno” dropped almost 33 percent, from $30 million to less than $20 million. Meanwhile, the rating is 8.0 (396,649 users) for “District 9” but 6.7 (94,848 users) for “Bruno” (www.imdb.com). This shows that the spread of information through social learning is fairly quick, immediately after the opening week, and exerts a huge impact on a movie’s later box office performance. This motivates me to establish a simplified two-period model to analyze consumer learning and firms’ optimal decisions about advertising, which helps to ease estimation computation while still capturing the main information features of this market.

The dataset used in this analysis covers movies that were widely released in U.S. theaters from 2000 through 2005. The dataset includes only movies that opened in more than 600 theaters and excludes “limited release” movies. Widely-released movies are considered national releases and, as such, require mass media advertising. To control the information spillover effect of movie sequels, I focus only on the first movie of a series. As a result, 632 out of 849 movies are included in the dataset. Data about observed movie characteristics as well as weekly market shares come from online sources (boxofficemojo.com, imdb.com, Yahoo Movie, etc.).

---

9 Limited release means that the studio first releases the movie in a small number of theaters and then expands to a large number of theaters if the movie performs well in box office. Wide release means that the studio releases the movie nationwide from the very beginning.

10 I developed an excel application using VBA (Visual Basic for Application) that can connect the
theater lifetime is collected and provided by TNS Media Intelligence. I have weekly advertising spending for each movie in my sample across media including broadcast, cable TV, newspapers, outdoor billboards, magazines, radio, and internet. To fit my simplified two-period model of studios’ optimal advertising spending decisions, I aggregate the advertising spending across media and divide the total theatrical advertising spending into two categories: pre-release advertising and post-release advertising.

My dataset also includes several important observable characteristics of the movies. Those include production budget, season indicators, Motion Picture Association of America (MPAA) ratings, genres, distributors, critic ratings, number of competitors, and runtime. Since a large proportion of a movie’s production budget covers salaries for stars, producers and directors, the cost of screenplay rights and cost of visual effects, therefore, production budget can be used as a proxy for star appeal, director appeal, story familiarity, and potential visual effects of a movie. In my sample, movie production budgets average around $44.60 million. Two season indicators, holiday and summer, account for the seasonality of the movie industry. The holiday indicator equals to 1 if the movie is released around Thanksgiving and Christmas, and the summer indicator equals to 1 if the movie is released between Memorial Day and Labor Day. In my sample, around 10% movies are released in the holiday season and around 28.5% movies are released in the summer season. There are four MPAA ratings including G, PG, PG-13, and R for movies in my sample. About 48.6% movies in my sample are rated as “PG-13” and 33% movies are rated as “R.” Only 3% of movies in the sample are rated as “G.” There are dozens of movie genres and sub-genres from which viewers can choose. However, several major genres make up the majority of popular movies. In my sample, most movies fall into those major genres for which I create five nonexclusive dummies: action, comedy, drama, family, and horror. About 40% of movies fall into the comedy genre, although they also can be categorized as both drama and action movies at the same time. Distributors are divided into major, mini-major, among others. Major distributors include Buena Vista, Fox, Miramax, Paramount, Sony, Warner Bros., and Universal. Mini-major distributors include DreamWorks, Lions Gate, and MGM. Those distributors make marketing and distributing decisions for about 90% of the movies in the U.S. market. Critic reviews inform moviegoers a movie’s quality before they actually watch the movie. Its value ranges from 0 to 100 and the average critic review score is 45 in my sample. The average runtime of a movie is 105.07 minutes. In this paper, I assume movie studios of new movies play a competitive monopoly game; therefore, they make decisions for each movie independently without taking their rivals’ reaction into account. Still, I include the number of other movies released widely in the same week to control for competitive

website and download the data automatically. All the downloaded data were stored to a local Access database.

11My data sample has no movie rated as “NC-17” which means “no one 17 and under admitted,” because NC-17 movies are usually limited released.
effects. In the sample, there are 2.32 other movies released in the same week, on average. Table 1 provides detailed descriptive statistics of those main variables in my dataset.

The market size is the number of U.S. population reported by the Census Bureau in a given period. Market shares are box office ticket sales of each movie divided by market size. The outside good market share is one minus the share of the movie.

4 The Model

In this section, I set up a generalized model to focus on the equilibrium advertising strategy of the studio and the evolution of consumer belief. The model can be broken into five components: (1) model primitives, (2) the information structure of the market, (3) demand, (4) supply, and (5) the pure strategy Nash equilibrium. I will discuss each part in turn.

4.1 Primitives

4.1.1 Players

There is a single studio with a new movie. The studio’s payoff depends on the expected total box office revenue it can collect from the movie. To maximize its payoff, the studio chooses its advertising spending, taking ticket price as given.

Consumers learn about the arrival of a new movie through advertising and then make movie-watching decisions. The quality (entertainment value) of a new movie is not fully observable prior to consumption, so consumers make their consumption decision based on the expected quality of a new movie.

4.1.2 Timing

The introduction of a new movie is modeled as an extensive form game, and Figure 3 shows the timing of the game. Time is divided into three periods.

Period 0 (after a movie is produced): a single studio produces a new movie \( j \), with observable attributes \( x_j \) and unobservable quality \( q_j \). Instead of knowing \( q_j \) perfectly, the studio receives a noisy signal of its movie’s quality, \( q_{js} \). Then the studio decides the optimal advertising spending for period 1 and period 2.

Period 1 (pre-release weeks and opening week of a movie): The new movie is introduced by advertising to the market. After being informed by studio advertising and

\[^{12}\]I also try two alternative definitions of market size: 66% and 33% of U.S. population. The estimation results don’t change much. The first definition approximates the proportion of people in U.S. who watch the movie at least once during a year and the second one approximates the regular moviegoer segment in U.S. who watch at least 6 movies a year in cinemas.
getting chances to enter the market, consumers update their beliefs about the movie’s quality with new information and decide whether to watch the movie in the opening week. At the end of period 1, some consumers may pass the information about the movie’s quality to potential consumers for period 2. Also, the studio updates its belief about its own movie’s quality and adjusts its advertising spending for period 2.

**Period 2 (post-release weeks of a movie):** informed consumers receive WOM information, take it as a noisy signal of the movie’s true quality, update their beliefs, and make their consumption decisions. Then the game ends.

### 4.2 The information structure of the market

In this section, I discuss the information structure of the theatrical market for new movies in details. More specifically, I will discuss the information learning processes of both supply and demand sides in this market, the roles of advertising and WOM playing during those learning processes and the interaction between advertising and WOM. When a new movie is released in the theater, it has both observed attributes and unobserved quality. Both the studio and consumers learn about the unobserved true quality of the movie through different information sources. I assume that the true unobserved quality of a new movie, \( q_j \), is a random draw from its population distribution \( \bar{q} \sim iidN (\bar{q}, \sigma_q^2) \).

#### 4.2.1 Information Learning: Studio

As discussed in section 3, the studio of a new movie can conduct various up-front assessments such as test screening and tracking surveys to learn about its movie’s potential playability and marketability. Therefore, for the model, I assume that the movie studio receives a noisy signal, \( q_{js} \), of the movie’s true quality \( q_j \) in period 0. \( q_{js} = q_j + \varepsilon_{js} \), with \( \varepsilon_{js} \sim iidN (0, \sigma_{\varepsilon}^2) \), is known only by the studio. Here, \( \sigma_q^2 \) measures how accurately the studio can learn about its movie’s true quality through up-front assessments. I assume that the studio uses the information from \( q_{js} \) to update its prior expectation of \( q_j \) according to the Bayesian updating rule:

\[
E_{s1}^s(q_j) = E^s[q_j | q_{js}] = \bar{q} + \beta_q(0) (q_{js} - \bar{q})
\]

where \( \beta_q(0) = \frac{\sigma_q^2(0)}{\sigma_q^2(0) + \sigma_s^2} = \frac{\sigma_q^2}{\sigma_q^2 + \sigma_s^2} \) is the weight the studio puts on \( q_{js} \). \( E_{j1}^s(q_j) \) is the weighted average of prior expected value of \( q_j \) and the noisy signal \( q_{js} \). When the signal is more accurate (with smaller value of \( \beta_q^s \)), more weight should be put on the signal received by the studio. The perception variance by the studio for period 1 is given by

\[
\sigma_q^2(1) = \frac{1}{\sigma_q^2(0) + \frac{1}{\sigma_s^2}} = \frac{\sigma_q^2}{\sigma_q^2 + \sigma_s^2} = \sigma_q^2 \left( \frac{\sigma_s^2}{\sigma_q^2 + \sigma_s^2} \right)
\]
This equation suggests that the perceived variance by the studio is lower than the prior variance perceived by consumers, before any extra information available to consumers. Since distribution \( f(q_{js} \mid q_{j}) \) satisfies the Monotone Likelihood Ratio Property (MLRP) in \( q_{j} \), for any value \( q^{*} \), \( \Pr(q_{j} \geq q^{*} \mid q_{js}) \geq (q_{j} \geq q^{*} \mid q_{js}) \) if \( q_{js} > q_{js} \). Intuitively, the better the received signal \( q_{js} \) is, the probability that the movie has quality above certain level is higher and the value of \( E_{js}^{\pi}(q_{j}) \) is higher. The studio then decides its optimal advertising spending \( a_{j1} \) according to its perceived movie quality \( E_{js}^{\pi}(q_{j}) \).

After the opening weekend, studios collect more information about their movies to adjust their advertising spending. For simplicity, I assume that the movie’s true quality \( q_{j} \) is revealed to the studio at the end of Period 1 and the studio adjusts its optimal advertising spending \( a_{j2} \) for period 2 according to \( q_{j} \).

### 4.2.2 Information Learning: Consumers

In this section, I discuss how advertising and WOM, as two main information channels, impact consumers’ information learning in the movie theatrical market, and how those two information channels interact with each other.

**The Role of Advertising** On the demand side, consumers need to learn two types of information: first, consumers need to know there is a new movie coming as well as its observed attributes; second, more importantly, consumers are motivated to learn about the movie’s unobserved quality. Both types of information can be carried by advertising. After the studio decides its advertising spending, advertisements are sent out as a series of messages, as we can see on TV, in theater or in mailbox. Consumers observe the advertisement intensity/frequency. When a consumer receives at least one advertisement, she is reached by the studio and knows that this new movie is coming to the theater. In this case, advertising impacts the demand by providing direct information about the movie and plays its “reaching role”. Also the consumer may use advertisement intensity to infer the movie’s quality, updates her belief, and then decides whether to watch the movie or not. In this case, advertising indirectly shows the studio’s confidence on the movie and plays its “signaling role”.

Suppose that, at the beginning of period 1, the optimal advertising spending the studio decides is \( a_{j1} \), and then advertisements are sent out to consumers as a series of messages. A consumer can receive 0, 1, 2, … advertisements. Following Butters (1977), I assume that the seller drops its advertisements at random into buyers’ “mailboxes.”

Therefore, the probability that consumer \( i \) receives \( k_{ij1} \) advertisements in period 1 is

---

13This process can be described by the classical urn model of probability theory: the mailboxes correspond to the \( n \) “urns,” and the advertisements correspond to the \( r \) “balls,” which are assumed to be assigned independently with equal probability to each of the urns. Therefore, for any given \( r \) and \( n \), the probability that a buyer receives \( k \) advertisements is given by the binomial distribution which approaches the Poisson distribution as \( n \) get large, holding \( \frac{r}{n} \) fixed.
given by the binomial distribution which approaches the Poisson distribution. \( k_{ij1} \) is the realized advertisement intensity observed by consumer \( i \) from \( \tilde{k}_{ij1} \sim pois(\lambda a_{j1}) \). Here, \( \lambda \) is the “reaching efficiency” parameter which can be used to quantify how efficiently advertisements can reach the market. I assume that consumers can be informed of the arrival of a new movie only by receiving advertisements, so the probability that consumer \( i \) is informed about the new movie is

\[
\varphi_1 (a_{j1}) = \text{prob} (k_{ij1} > 0) = 1 - \text{prob} (k_{ij1} = 0) \\
= 1 - \exp (-\lambda a_{j1})
\]

which is also the market coverage rate.

In the post-release weeks, the studio adjusts its optimal advertising spending to \( a_{j2} \), and additional advertisements are sent out to reach or remind more potential consumers. Consumers who are aware of the new movie comprise two groups: consumers who are informed by advertisements in period 1 and still remember the movie, and consumers who are reminded or just informed by advertisements in period 2. The proportion of the covered market or the probability that consumer \( i \) is aware of the new movie can be written as:

\[
\varphi_2 (a_{j1}, a_{j2}) = \varphi_1 (\mu a_{j1}) + (1 - \varphi_1 (\mu a_{j1})) \text{prob} (k_{ij2} > 0) \\
= [1 - \exp (-\lambda (\mu a_{j1} + a_{j2}))]
\]

Here, \( \mu \) describes how effectively advertising money still works in period 2, therefore, \((1 - \mu)\) describes the depreciation rate of advertising stock because of consumers’ memory loss over time.

Consumers fully learn from all available information and infer product quality through the studio’s actions. In my model, advertising is not only used to inform consumers of the availability of a new movie, but also allowed to be used by consumers to infer the movie’s quality. With the existence of a signaling equilibrium, the studio’s advertising spending in period 1 \( a_{j1} = A_1 (q_{js}) \) is an increasing function of received noisy signal, \( q_{js} \). Intuitively, when the studio receives a better signal \( q_{js} \), it spends more on advertising, and consumers tend to observe higher advertising intensity \( k_{ij1} \). Therefore, a rational consumer should take \( k_{ij1} \) as a noisy signal of \( a_{j1} \) and update her belief about the movie’s quality. According to the Bayesian updating rule, the posterior distribution and expectation of the true quality \( q_j \) for consumer \( i \) in period 1 after observing advertisement density \( k_{ij1} \) are

\[
g_{i1} (q_j | k_{ij1}, A_1 (q_{js})) = \frac{\int \int f (k_{ij1} | A_1 (q_{js})) f (q_{js} | q_j) d_{q_{js}} d_{q_j} g_0 (q_j)}{\int \int f (k_{ij1} | A_1 (q_{js})) f (q_{js} | q_j) d_{q_{js}} d_{q_j} g_0 (q_j) d_{q_j}}
\]
The Role of WOM  

After the opening week, some consumers who already watched the movie in period 1 may pass the information about the movie’s true quality to potential consumers who enter the market in period 2. When consumer \( h \) talks to consumer \( i \) about the movie’s quality, consumer \( i \) receives a WOM signal \( q_{ijw} \) which is a random drawn from distribution \( iidN (q_j, \sigma^2_w) \), where \( q_j \) is the true quality of the movie and \( \sigma^2_w \) is the variance of the WOM signal. Consumer \( i \) may get several WOM signals and aggregate all information. Let \( d_{j1} \) be the number of tickets sold in period 1, which is also the movie’s box office performance in the opening week. \( \rho \) is the average proportion of consumers who like to share their movie-watching experiences with the representative consumer \( i \). Here, \( \rho \) measures the information transmission speed, and the higher value of \( \rho \) implies that more consumers like to spread information through WOM. Therefore, consumer \( i \) gets a sample mean of experience signals, \( \bar{q}_{ijw} \), which is a random drawn from distribution \( iidN \left( q_j, \frac{\sigma^2_w}{\rho d_{j1}} \right) \). The variance of \( \bar{q}_{ijw} \) is a decreasing function of \( d_{j1} \), which simply means that if more consumers watch the movie in period 1, the average WOM signal, \( \bar{q}_{ijw} \), is more accurate about the movie’s true quality, \( q_j \).

Besides WOM signals, there is another information source about the movie’s true quality available to consumers in period 2: the movie’s box office performance, \( d_{j1} \), in period 1. With the existence of a signaling equilibrium, \( d_{j1} = D_1 (A_1 (q_{js})) \) should be an increasing function of \( q_{js} \) and a rational consumer should infer \( q_{js} \) from \( d_{j1} \). Therefore consumer \( i \) may take those two information sources into account and updates her beliefs according to the Bayesian rule. The posterior distribution and expectation of the quality \( q_j \) for consumer \( i \) in period 2 is

\[
E^c_{ij2} \left[ q_j \mid \bar{q}_{ijw}, d_{j1}, A_1 (q_{js}) \right] = \int q_j \cdot g_{i2} \left( q_j \mid \bar{q}_{ijw}, d_{j1}, A_1 (q_{js}) \right) dq_j
\]

(2)

If consumer \( i \) infers \( q_{js} \) from \( d_{j1} \), then the prior expected \( q_j \) for consumer \( i \) is \( E^c_{ij1} \left[ q_j \mid d_{j1} \right] = E^c_{j1} (q_j) = \bar{q} + \beta_q (0) (D_1^{-1} (d_{j1}) - \bar{q}) \), and equation (2) can be written as simple as

\[
E^c_{ij2} (q_j) = E^c_{ij1} \left[ q_j \mid d_{j1} \right] + \beta^c_q (1) (\bar{q}_{ijw} - E^c_{ij1} \left[ q_j \mid d_{j1} \right])
\]

where \( \beta^c_q (1) = \frac{\sigma^2_q (1)}{\sigma^2_q (1) + \frac{\sigma^2_w}{\rho d_{j1}}} = \frac{\sigma^2_q (1)}{\sigma^2_q (1) + \frac{\sigma^2_w}{\rho d_{j1}}} \) is the weight that consumer \( i \) put on the WOM
signal $q_{ijw}$. The perception variance by consumer $i$ for period 2 is

$$\sigma_q^2(2) = \frac{1}{\sigma_q^2(1) + \rho d_{j1}^2} = \frac{1}{\sigma_q^2(1) + \rho d_{j1}^2} = \frac{\sigma_w^2}{\rho d_{j1}^2} \beta_q^c (1)$$

Here, I discuss the intuitions of how advertising and WOM interact with each other and how the interaction between them supports the signaling role of advertising. On one hand, WOM between consumers influences the long-term return to advertising. Imagine that if a studio with a bad movie spends a lot on advertising to pretend having a good one, the WOM after the opening week reveals the true quality and fewer consumers will choose to watch the movie in the post-release weeks. Therefore, low-quality movie would not recover such an expensive investment. This asymmetry in the returns to advertising created by WOM forces firms to decide its advertising spending according to its movie’s quality, therefore advertising can be a credible quality signal. On the other hand, advertising influences WOM process as well. When studios decide their pre-release advertising spending, they need to understand that the pre-release advertising not only impacts how many people will watch the movie in the opening week, but also it indirectly impacts how many people will talk about the movie after the opening week. The more consumers are induced by advertising to watch the movie in the opening week, the more WOM communications happen in the post-release weeks. Then the information revealed to consumers is more accurate. This mechanism further prevents the studios from aggressively advertising a movie if they think the movie is a bad one.

4.3 Demand

On the demand side, consumers make static discrete choice about whether to watch movie $j$ conditional on they are reached by advertising. Consumer $i$’s expected utility from watching movie $j$ in period $t$ $(t = 1, 2)$:

$$E u_{ijt} = \gamma x_j + \alpha T D_t + E_{ijt}^c [q_j | I_i (t)] + \eta_{jyt} + \varepsilon_{ijt}$$

Here, $x_j$ is composed of observed characteristics of movie $j$, such as genre, production budget, studio, the MPAA rating, the holiday indicator, etc., and $\gamma$ is composed of consumer taste parameters. $T D_t$ is the time dummy which indicates whether it is period 2 or not and $\alpha$ is the utility weight that consumer $i$ attaches to $T D_t$. $E_{ijt}^c [q_j | I_i (t)]$ is the expected quality of movie $j$ perceived by consumer $i$ conditional on her information set, $I_i (t)$ in period $t$, and $\eta_{jyt}$ is the realized aggregate demand shock in period $t$ from $\tilde{\eta}_{jyt} \sim iid N \left(0, \sigma_{\eta_j}^2 \right)$, and $\varepsilon_{ijt}$ is consumer $i$’s realized idiosyncratic preference shock in period $t$ from $\tilde{\varepsilon}_{ijt} \sim iid EV$. Note price of watching a movie in the theater is not included in the utility function, since it is the same for movies of different quality levels and keeps
stable during the sample period. Consumer $i$’s utility from the outside option in week $t$ is $u_{it0} = \varepsilon_{it0}$ with mean utility normalized to zero, and $\varepsilon_{it0} \sim iidEV$.

Consumers are assumed to be myopic in the sense that they do not make decisions intertemporally. Therefore consumer $i$’s watching decision in period 1 is described as:

$$w_{ij1} = w_{i1}(k_{ij1}; A_1(q_{js})) = \begin{cases} 1 & \text{if } Eu_{ij1} \geq 0 \\ 0 & \text{if } Eu_{ij1} < 0 \end{cases},$$

and her watching decision in period 2 is described as:

$$w_{ij2} = w_{i2}(\bar{q}_{ijw}, d_{j1}; A_1(q_{js})) = \begin{cases} 1 & \text{if } Eu_{ij2} \geq 0 \\ 0 & \text{if } Eu_{ij2} < 0 \end{cases}.$$ (4)

Both equations (3) and (4) tell us that an informed consumer $i$ will watch the movie $j$ in period $t$ only when $Eu_{ijt} = \gamma x_j + \alpha TD_t + E_{ijt}^c[q_j|I_t(t)] + \eta_{jt} + \varepsilon_{ijt} \geq \varepsilon_{it0}$. Then the probability that the informed consumer $i$ chooses to watch the movie $j$ in period $t$ is $\tau_{it}\left(\delta_{jt}, E_{ijt}^c[q_j|I_t(t)]\right) = \Pr\left(Eu_{ijt} > Eu_{it0}\right) = \frac{\exp(\delta_{jt} + E_{ijt}^c[q_j|I_t(t)])}{1 + \exp(\delta_{jt} + E_{ijt}^c[q_j|I_t(t)])}$, where $\delta_{jt} = \gamma x_j + \alpha TD_t + \eta_{jt}$ is the “mean utility” for movie $j$. In period 1, the probability that consumer $i$ chooses to watch the movie, conditional on observing $k_{ij1} > 0$ is given by

$$\tau_{ij1} = \frac{\exp(\delta_{j1} + E_{ij1}^c[q_j|k_{ij1}, A_1(q_{js})])}{1 + \exp(\delta_{j1} + E_{ij1}^c[q_j|k_{ij1}, A_1(q_{js})])} = \tau_{ij1}(k_{ij1}, q_{j1}; x_j, A_1(q_{js}))$$

With advertising spending $a_{j1}$, the market share of movie $j$ in period 1 is:

$$s_{j1} = \varphi_1(a_{j1}) \sum_{k_{ij1}=1}^{\infty} \tau_{ij1}(k_{ij1}, q_{j1}; x_j, A_1(q_{js})) f(k_{ij1}|a_{j1})$$

$$= \varphi_1(a_{j1}) \tau_1(a_{j1}, q_{j1}; x_j, A_1(q_{js}))$$

$$= S_1(a_{j1}, q_{j1}; x_j, A_1(q_{js}))$$

and, with market size $M$, the number of ticket sold in period 1 is $d_{j1} = s_{j1}M$.

In period 2, the probability that consumer $i$ chooses to watch movie $j$, conditional on her being reached by advertisements is given by

$$\tau_{ij2} = \frac{\exp(\delta_{j2} + E_{ij2}^c[q_j|\bar{q}_{ijw}, d_{j1}, A_1(q_{js})])}{1 + \exp(\delta_{j2} + E_{ij2}^c[q_j|\bar{q}_{ijw}, d_{j1}, A_1(q_{js})])} = \tau_{ij2}(\bar{q}_{ijw}, d_{j1}, q_{j2}; x_j, A_1(q_{js}))$$
With advertising spending \( a_{j2} \), the market share of movie \( j \) in period 2 is:

\[
s_{j2} = \left[ \varphi_2 (a_{j1}, a_{j2}) - s_{j1} \right] \int \tau_{ij2} \left( \tilde{q}_{ijw}, d_{j1}, \eta_{j2}; x_j, A_1 (q_{js}) \right) \phi \left( \tilde{q}_{ijw} \mid q_j; \frac{\sigma^2_w}{\rho d_{j1}} \right) d\tilde{q}_{ijw} \tag{6}
\]

\[
= \left[ \varphi_2 (a_{j1}, a_{j2}) - s_{j1} \right] \tau_{j2} \left( q_j, \eta_{j2}; d_{j1}, x_j, A_1 (q_{js}) \right) 
\]

\[
= S_2 \left( a_{j2}, q_j, \eta_{j2}; a_{j1}, d_{j1}, x_j, A_1 (q_{js}) \right)
\]

where \( \left[ \varphi_2 (a_{j1}, a_{j2}) - s_{j1} \right] \) is the proportion of potential consumers who are aware of the new movie but haven’t watched the movie yet. From equation (5) and (6), we can tell that the market share of the new movie in each period is composed of two parts: the proportion of consumers who are reached by advertising and haven’t watched the movie yet, and the proportion of consumers who are convinced to watch the movie.

### 4.4 Supply

The supply side is modeled as a monopolistic competition problem. Unlike other product markets, studios make decisions about optimal advertising spending when releasing new movies, instead of choosing an optimal price. For each movie, the studio makes decisions independently, taking its rivals’ actions as given. The studio chooses optimal advertising spending \( a_{jt} \) in period \( t = 1, 2 \), based on its information about the movie’s quality. I denote \( SV_{jt} = \{ E_{jt} (q_j), \sigma^2_q (t), x_j, \delta_{j(t-1)}, a_{j(t-1)} \} \), as the set of state variables that are relevant to the decision of the studio. The per period expected profit for the studio of movie \( j \) in period \( t = 1, 2 \) is

\[
\pi_{jt} (a_{jt}; SV_{jt}) = \int \left[ S_t \left( \eta_{jt}; SV_{jt}, a_{jt} \right) M_p - a_{jt} \right] d\Phi(\eta_{jt})
\]

Then the value function for the studio is

\[
V (SV_{j1}) = \max_{a_{j1} \geq 0} \left[ \pi_{j1} (a_{j1}; SV_{j1}) + E_{j1} V (SV_{j2} | a_{j1}, SV_{j1}) \right]
\]

\[
V (SV_{j2}) = \max_{a_{j2} \geq 0} \left[ \pi_{j2} (a_{j2}; SV_{j2}) \right]
\]

where \( E_{j1} V (SV_{j2} | a_{j1}, SV_{j1}) = \int V (q_j | a_{j1}, d_{j1}, SV_{j1}) d\Phi(q_j | I(1)) \) and \( q_j | I(1) \) follows normal distribution with mean \( E^c_{ij2} (q_j) \) and variance \( \sigma^2_q (1) \). It should be noted that the studio explicitly takes into account the effect of its advertising decision \( a_{j1} \) on the next period’s expected mean quality \( E^c_{ij2} (q_j) \) and variance \( \sigma^2_q (2) \) perceived by consumers through opening week market performance \( d_{j1} \).

By solving above profit maximization problems, we have the optimal advertising
spending for period 2 as

\[ a^*_j = \max \left( \frac{\ln (\lambda M p E_2 \left( \tau_{j2}(q, \eta_{j2}; d_{j1}, x_j, A_1(q_{js})) \right))}{\chi} - \mu a_{j1}, 0 \right) \]  

(7)

where

\[ E_2 \left( \tau_{j2}(q, \eta_{j2}; d_{j1}, x_j, A_1(q_{js})) \right) = \int \tau_{j2}(q, \eta_{j2}; d_{j1}, x_j, A_1(q_{js})) d_{\Phi(\eta_{j2})} = \tau_{j2}(q_j d_{j1}, x_j, A_1(q_{js})). \]

The optimal advertising spending for period 1, \( a^*_{j1} = A_1(q_{js}; x_j) \), satisfies the equilibrium condition

\[ \frac{\partial \left[ \pi_{j1}(a_{j1}; SV_{j1}) + E_{j1}V(SV_{j2}|a_{j1}, SV_{j1}) \right]}{\partial a_{j1}} \bigg|_{a_{j1}=a^*_{j1}} = 0 \]

(8)

Here, I assume that studios maximize the total expected profit from the theatrical market by choosing the optimal advertising expenses, without considering the complexity of the vertical structure in this market. There are three key stages in the value chain in the theatrical movie market: production, distribution, and exhibition. Each stage involves different types of entities such as major studios, independent production companies, independent distributors, national exhibition chains and regional exhibitors. Vertically integrated major studios are often simultaneously engaged in both production and distribution, as well as interacting with exhibitors. In practice, movie studios pay the full expense for national marketing, but movie studios and exhibitors split the movie box office revenue according to the contractual arrangements between them. The general rule is that the distributor’s share is high in the first few weeks, and it declines as the movie’s run proceeds (Vogel 2001). Ideally, it is better to incorporate the optimal decisions of both distributors and exhibitors as well as considering the impact of the contractual agreements between them. However, I simplify the model by ignoring the contractual complexity between different entities for the following reasons: first, the movie’s box office performance positively impacts the revenue from other nontheatrical windows. Second, distributors and exhibitors normally have a long-term relationship for many movies and they have multiple negotiating points such as the length of the run in the theater and the number of screens the movie can be promised. Therefore it might be in distributors’ best interest to consider exhibitors’ interest when making advertising decisions. Third, the simplification keeps the model trackable but still rich enough to investigate the questions in interest.

4.5 Advertising-Watching Equilibrium

Since this paper mainly investigates the empirical implications of how studios use advertising to manipulate sales in a learning environment, so I will focus on discussing the existence of pure strategy separating Nash equilibrium of this incomplete information
game in this section. In equilibrium, both demand and supply sides have rational expectation about each other’s strategies and all expectations are consistent with the actual strategies.

**Definition 1** The rule $A_t(\cdot)$ and $w_{it}(\cdot)$ constitute an equilibrium provided each is a best response to the other. That is, $(A_t(\cdot), w_{it}(\cdot))$ is an equilibrium if

(E1) $A_1(\cdot) \in \arg \max [\pi_1(a_{j1}; SV_{j1}) + E_1V(SV_{j2}|a_{j1}, SV_{j1})]$

and $A_2(\cdot) \in \arg \max [\pi_{j2}(a_{j2}; SV_{j2})]$

(E2) $w_{it}(\cdot) = 1$ if and only if $\delta_{jt} + E^c_{ijt}[q_j|I_i(t)] + \varepsilon_{ijt} \geq \varepsilon_{iwt}$.

To discuss the existence of a pure strategy Nash signaling equilibrium, we discuss the following lemmas first.

**Lemma 2** If the advertising policy function $A_1(q_{js}, x_j)$ is increasing in $q_{js}$, then $E_{uij1} = \delta_{j1} + E^c_{ij1}[q_j|k_{ij1}, A_1(q_{js})] + \varepsilon_{ij1}$ is increasing in $k_{ij1}$, and the best response rule is

$$w_{i1}(k_{ij1}; A_1(q_{js})) = \begin{cases} 1 & \text{if } k_{ij1} \geq \left[k_{ij1}^*\right] \\ 0 & \text{if } k_{ij1} < \left[k_{ij1}^*\right] \end{cases}$$

where $k_{ij1}^*$ is defined by $\delta_{j1} + E^c_{ij1}[q_j|I_1(1)] + \varepsilon_{ij1} = \varepsilon_{i01}$, and $[k_{ij1}^*]$ is the smallest integral which is not smaller than $k_{ij1}^*$. Therefore the proportion of consumers convinced to watch movie $j$, $\tau_1(a_{j1}, \eta_{j1}; x_j, A_1(q_{js}))$, is an increasing function of $a_{j1}$ conditional on $x_j$.

**Proof.** Since $k_{ij1} \sim \text{pois}(\lambda a_{j1})$ and the family of Poisson distributions satisfy MLRP, the posterior CDF of $a_{j1}$, $H(a_{j1}|k_{ij1})$, is a decreasing function of $k_{ij1}$. Since $A_1(\cdot)$ is an increasing smooth function of $q_{js}$, the posterior CDF of $q_j$, $G(q_j|k_{ij1})$, is also a decreasing function in $k_{ij1}$. If $k_{ij1} > k_{hj1}$, then $\Pr[q_j > \tilde{q}|k_{ij1}] > \Pr[q_j > \tilde{q}|k_{hj1}]$ for any value $\tilde{q}$, so $q_j|k_{ij1}$ first-order stochastically dominates $q_j|k_{hj1}$ and $E[q_j|k_{ij1}] > E[q_j|k_{hj1}]$. Therefore $E_{uij1}$ increases in $k_{ij1}$ and $\tau_1(a_{j1}, \eta_{j1}; x_j, A_1(q_{js}))$ is an increasing function of $a_{j1}$. 

**Lemma 3** $E_{uij2} = \delta_{j2} + E^c_{ij2}[q_j|\overline{q}_{ijw}, d_{j1}, A_1(q_{js})] + \varepsilon_{ij2}$ is increasing in $\overline{q}_{ijw}$ and $\tau_2(q_j, \eta_{j2}; d_{j1}, x_j, A_1(q_{js}))$ is increasing in $q_j$.

**Proof.** Above lemma only requires $E^c_{ij2}[q_j|\overline{q}_{ijw}, d_{j1}, A_1(q_{js})]$ increases in $\overline{q}_{ijw}$. The assumption that $\overline{q}_{ijw} \sim \text{iid}N(q, \sigma^2_{pd_{j1}})$ ensures this lemma holds.$^{14}$

**Assumption A1**: $s_{j1}$ is small enough.

$A1$ means consumers have idiosyncratic preference shocks for the movie because of outside options, and the proportion of consumers who are informed by advertisements and choose to watch the movie in period 1 is small. Intuitively, this assumption requires that the market in long-run should be important enough for a movie’s success. This should be

$^{14}$As long as the family of distributions $f\left(\overline{q}_{ijw} | q_j, \sigma^2_{pd_{j1}}\right)$ has MLRP in $q_j$, then this lemma holds.
a very reasonable assumption with supporting evidence from the data collected. Table 2
shows that the market share in the opening week for movies in my data sample is 0.82%
on average, with maximum value equaling to 4.18%. While, the market share in the
post-release weeks is around 2% on average, with maximum value being more than 15%.

With the increasing impact of WOM on movies’ box office performance, one may
argue that movie studios with low quality movies may strategically spend more on pre-
release advertising. By focusing on the short-run market performance, those studios can
recoup their investments before any negative WOM generated. To show that it is hardly
the case, I compare movies for which more than 95% advertising budget was used in
the pre-release weeks to movies for which less than 60% advertising budget was used for
the same period. From Table 2, we can see that the first group of movies does collect
almost half of their total box office revenue from the opening weekend on average, while
the second group of movies mainly depends on the long-run box office performance. The
average online critic rating and moviegoer rating for the first group of movies are much
lower than those for the second group of movies. However, on average, the first group
of movies can only recover around 63% of their advertising investment by collecting box
office revenue. In contrast, the second group of movies collects 390% of their advertising
spending through box office revenue on average. Figure 4 further shows that the ratio of
pre-release advertising spending to the total advertising spending and the ratio of opening
weekend box office revenue to the total box office revenue are negatively correlated to the
profitability of the movie which is shown by the ratio of total box office revenue to the
total advertising spending. Therefore, it is not rational for movies studios to focus only
on short-run market by aggressively advertising in pre-release advertising and ignore the
negative impact of bad WOM on the long-run market.

**Assumption A2**: \( \frac{\tau_{12}(q_j, q_{2d_1}, x_j, A_1(q_{js}))}{d_{j1}q_{ij}} \) is nonnegative or limited negative.

In period 2, consumers have two information sources from which to update their
beliefs: the movie’s market performance in period 1 and the WOM among consumers
about the movie. \( d_{j1} \) is determined by the studio’s advertising action and, therefore, can
be called as “firm-generated” information. \( q_{ijw} \) is determined by WOM communication
among consumers and, therefore, can be called “consumer-generated” information. A2 implies
that two types of information are primarily complements, or, if they are substi-
tutes, the ratio is small enough.

**Lemma 4** If \( A_1(\cdot) \) is a best response to \( w_{id}(\cdot) \), then with assumptions A1 and A2, \( A_1(\cdot) \)
is nondecreasing, and, for \( q_{js} \in Q_S \subseteq Q_{S15} \), \( A_1(\cdot) \) is increasing.

**Proof.** In order to prove the result, we apply Theorem 2 in Athey (2002). To verify
that our model meets all the requirement of Theorem 2 in Athey (2002), I rewrite the
\[15\] \( Q_s \) is the domain of random variable \( q_{js} \) and \( Q_{S15} \) is the subset of the domain of \( q_{js} \).
Increasing for
With assumptions A1 and A2, we know
that
\( f \)
Athey (2002), we know
is log-spm
than
\( k \)
(D1), (E2), and
De
\( \text{Definition 5} \)
check the sign of
\( q \)
js
studio’s revenue maximization problem in the following way:

\[
\max_{a_{j1}} \Pi (a_{j1}, q_{js}) = \int \pi (a_{j1}, q_j) g (q_{j} \mid q_{js}) dq_j
\]

s.t. \( \pi (a_{j1}, q_j) = \int \left[ s_{j1} M_p + (\varphi_2 (a_{j1}, a_{j2}) - s_{j1}) M_p \int \tau_2 (q_j, \eta_{j2}, d_{j1}) d_{\phi (\eta_{j2})} \right] d_{\phi (a_{j1})} \]

\( s_{j1} = S_1 (a_{j1}, \eta_{j1}) = \varphi_1 (a_{j1}) \tau_1 (a_{j1}, \eta_{j1}) \)

\( a_{j2} = \max \left( \frac{\ln \lambda M_p E_2 (\tau_j, \eta_{j2}, d_{j1}) - \mu a_{j1}, 0}{} \right) = A_2 (a_{j1}) \)

\( \Pi \geq 0 \)

\( a_{j1} \geq 0 \)

\( a_{j2} \geq 0 \)

Theorem 2 in Athey (2002) requires that \( \pi (a_{j1}, q_j) \) satisfies SC2 in \( (a_{j1}, q_j) \) and \( g (q_j \mid q_{js}) \)
is log-spm\footnote{log-supernormal is abbreviated to log-spn. Here, it means that \( q_{js} \) shifts the conditional distribution of \( q_j \) according to the monotone likelihood ratio property.} as a minimal pair of sufficient conditions for \( A_1 (q_{js}) \) to be nondecreasing in \( q_{js} \). \( g (q_j \mid q_{js}) \) is log-spn can be met by the assumption that \( g (q_j \mid q_{js}) \) is conditional normal distribution and has MLRP in \( q_{js} \). By assuming \( \pi (a_{j1}, q_j) \) is \( C^2 \), I just need to check the sign of \( \frac{\partial^2 \pi (a_{j1}, q_j)}{\partial a_{j1} \partial q_j} \).

\[
\frac{\partial^2 \pi (a_{j1}, q_j)}{\partial a_{j1} \partial q_j} = \left( \frac{\partial \varphi_2 (a_{j1}, a_{j2})}{\partial a_{j1}} - \frac{\partial s_{j1}}{\partial a_{j1}} \right) M_p \int \frac{\partial \tau_2 (q_j, \eta_{j2}, d_{j1})}{\partial q_j} d_{\phi (\eta_{j2})} \]

\( + (\varphi_2 (a_{j1}, a_{j2}) - s_{j1}) M_p \int \frac{\partial^2 \tau_2 (q_j, \eta_{j2}, d_{j1})}{\partial d_{j1} \partial q_j} \frac{\partial d_{j1}}{\partial a_{j1}} d_{\phi (\eta_{j2})} \)

With assumptions A1 and A2, we know \( \frac{\partial^2 \pi (a_{j1}, q_j)}{\partial a_{j1} \partial q_j} \geq 0 \). Then by using Theorem 2 in Athey (2002), we know \( A_1 (\cdot) \) is nondecreasing. With assumption A2 and the assumption that \( f (q_{js} \mid q_j) \) satisfies MLRP, \( A_1 (\cdot) \) cannot be constant for all \( q_{js} \in Q_S \), then \( A_1 (\cdot) \) is increasing for \( q_{js} \in Q_S' \subseteq Q_S \). ■

**Definition 5** A pure strategy Nash signaling equilibrium is an equilibrium which satisfies (E1), (E2), and

(E3) \( w_{i1} (k_{ij1}; A_1 (q_{js})) = \begin{cases} 
1 & \text{if} \quad k_{ij1} \geq \left[k^*_{ij1}\right] \\
0 & \text{if} \quad k_{ij1} < \left[k^*_{ij1}\right]
\end{cases} \) where \( k^*_{ij1} \) is defined by

\( \delta_{ij1} + E_{i1} [q_{ij} | I_i (1)] + \varepsilon_{i1} = \varepsilon_{i01} \), and \([k^*_{ij1}\] is the smallest integral which is not smaller than \( k^*_{ij1} \). Then \( \tau_1 (a_{j1}, \eta_{j1}; x_j, A_1 (q_{js})) \) is an increasing function of \( a_{j1} \).

(E4) \( A_1 (q_{js}) \) is nondecreasing in \( q_{js} \in S \), and for \( q_{js} \in S' \subseteq S \), \( A_1 (\cdot) \) is increasing.
4.6 Simplified Examples

To gain more intuition about the existence of the separating equilibrium of the model, I simplify the model in the following way. Instead of assuming that quality is a continuous variable, I assume it has only two possible values, either high (H) or low (L). In period 0, movie j’s quality, \( q_j \), is exogenously determined by nature, and consumers believe that it has probability \( \mu_0 \) to be \( q_H \) and \((1 - \mu_0)\) to be \( q_L \). The studio observes the quality signal \( q_{js} \) (either \( q_{HS} \) or \( q_{LS} \)), and the probability to be right is \( \eta \in (\frac{1}{2}, 1) \), which means \( \Pr(q_{HS}|q_H) = \Pr(q_{LS}|q_L) = \eta \) and \( \Pr(q_{HS}|q_L) = \Pr(q_{LS}|q_H) = 1 - \eta \). After receiving \( q_{js} \), the studio updates its belief of quality and decides the advertising spending \( a_{ij1} \). First, I assume that when the studio receives \( q_{HS} \), it will spend \( A_{H1} \); when it receives \( q_{LS} \), it will spend \( A_{L1} \). And \( A_{H1} > A_{L1} \) which will be proved to be the equilibrium result.

In period 1, consumers update their beliefs of movie j’s quality and make their consumption decisions. As in the general case, consumers cannot observe the advertising spending directly, but advertising intensity \( k_{ij1} \) drawn from \( \tilde{k}_{ij1} \sim \text{pois}(A_{ij1}) \) (\( j = H \) or \( L \)). The probability that \( q_j = q_H \) perceived by consumer \( i \) is updated to be

\[
\mu_1(k_{ij1}) = \frac{\text{prob}(q_H | k_{ij1})}{\text{prob}(q_H) + \text{prob}(q_L)}
\]

where \( \text{prob}(q_H | k_{ij1}) = \text{prob}(q_H | A_{H1}) \text{prob}(A_{H1}) + \text{prob}(q_H | A_{L1}) \text{prob}(A_{L1}) = \text{prob}(q_H | A_{H1}) \eta + \text{prob}(q_H | A_{L1}) (1 - \eta) \) and \( \text{prob}(q_L | k_{ij1}) = \text{prob}(q_L | A_{H1}) (1 - \eta) + \text{prob}(q_L | A_{L1}) \eta \). And it is easy to show that \( \mu_1(k_{ij1}) \) is an increasing function of \( k_{ij1} \). So consumer i’s expected utility\(^{17}\)

\[
E_{u_{ij1}} = E_{ij1}[q_j | I_i(1)] - p = (\mu_1(k_{ij1}) q_H + (1 - \mu_1(k_{ij1})) q_L) - p \geq 0
\]

determines the critical value of \( k_{ij1} : k_1^* \). If consumer \( i \) receives \( k_{ij1} \geq k_1^* \), she chooses to watch it, otherwise, she does not, as shown in Figure 5.

In period 2, the studio spends \( A_{j2} \) on advertising to reach consumers in period 2, and consumers update their beliefs of \( q_j \) based on information from WOM communication. Here, WOM is not a noisy signal for simplicity’s sake. Fraction \( \rho s_{j1} \) of consumers will know the true quality of the movie, and \((1 - \rho s_{j1})\) fraction of consumers will keep their prior perceived quality level \( \bar{q} = \mu_0 q_H + (1 - \mu_0) q_L \). WOM communication ratio, \( \rho \), is used to indicate the fraction of consumers who like to share their review of the movie with other consumers after watching it, and \( s_{j1} \) is the market share in period 1.\(^{18}\) Consumer

\(^{17}\)Here, consumers’ expected utility only depends on expected quality and price for simplicity’s sake.

\(^{18}\)Although WOM is not a noisy signal in this simple case, but assumptions here still make sure information about product quality is not revealed completely in period 2. And how many consumers have experienced the product in period 1 and the communication level still impact how well consumers in period 2 know about the product quality.
i’s utility in period 2 is

\[ E_{u_{ij2}} = E_{ij2}[q_j | I_i(2)] - p = q_j - p, \]

if she learns from WOM, \( j = H \) or \( L \)

\[ \bar{q} - p, \quad \text{otherwise} \]

Here, I assume \( \bar{q} - p < 0 \) and \( q_H - p > 0 \). Then the studio with \( q_j \) has market share in period 1 \( s_{j1} = \text{prob}(k_{ij1} > 0) \text{prob}(k_{ij1} \geq k^*_j | A_{j1}) \). In period 2, the studio with \( q_H \) has market share \( s_{j2} = \text{prob}(k_{ij2} > 0) \rho * \delta_{j1} \) and the studio with \( q_L \) has market share zero. However, when the studio makes advertising spending decisions in period 0, it is not completely sure about its movie’s quality as well as market share in period 2. So it is the expected market shares of period 2,

\[ E_0[s_{H2}|q_{Hs}] = \text{prob}(q_H|q_{Hs}) * s_{H2} \]

and

\[ E_0[s_{L2}|q_{Ls}] = \text{prob}(q_H|q_{Ls}) * s_{H2} \]

are used instead of realized market shares of period 2 for studios’ profit maximization problem.

I assume the whole market size is \( N_1 \) in period 1 and \( N_2 \) in period 2. The profit maximization problem for the studio receiving \( q_{Hs} \) is:

\[ \max_{A_{H1}, A_{H2}} \Pi_H = s_{H1}(A_{H1}) N_1 p + E_0[s_{H2}(A_{H1}, A_{H2}) | q_{Hs}] N_2 p - A_{H1} - A_{H2} \]

The profit maximization problem for the studio receiving \( q_{Ls} \) is:

\[ \max_{A_{L1}, A_{L2}} \Pi_L = s_{L1}(A_{L1}) N_1 p + E_0[s_{L2}(A_{L1}, A_{L2}) | q_{Ls}] N_2 p - A_{L1} - A_{L2} \]

Then at the end of period 1, studios update their beliefs about quality and adjust \( A_{j2} \), and only the studio with the high-quality movie will have advertising spending in period 2.

After solving the maximization problem, we can get \( A^*_{H1} > A^*_{L1} \). With proper parameter values, we have a separating equilibrium. Compared to the case in which consumers have complete information and know the true quality before watching the movie, we can have \( A^*_{H1} > A^*_H \) and \( A^*_{L1} > A^*_L \); compared to the case in which quality is uncertain, and there is no WOM in period 2 (\( \rho = 0 \)), we see there is no market for both types of movies. There is also no market in the case where quality is uncertain, and consumers are naive and do not use advertising spending to infer the quality level.

From this simple case, we can learn that WOM brings an additional benefit for the
firm with a high-quality product. This mechanism provides motivation for firms to signal their high-quality by advertising and assures the existence of a separating equilibrium. The positive information externality effect between consumers in both periods (consumers in period 1 reveal direct information to consumers in period 2 through WOM; consumers’ purchase in period 2 indirectly constrains firms’ advertising behavior, which indirectly reveals information to consumers in period 1 through advertising) and even helps the market to exist (for the other three cases, there is no market).]

5 Estimation Strategy

5.1 Likelihood Contribution

From data, I observe box office performance and advertising spending in two periods for movie \( j \): \( Y_j = (s_{j1}, s_{j2}, a_{j1}, a_{j2})' \). From equations (5), (6), (7) and (8), we know that the observed variables \( Y_j \) can be expressed as functions of unobserved random variables \( Z_j = (\eta_{j1}, \eta_{j2}, q_{js}, q_j)' \) in a more compact form. The relationship between \( Y_j \) and \( Z_j \) can be devoted to be \( Y_j = \zeta (Z_j|x_j, \Theta) \) which is also the Bayesian-Nash equilibrium equation. So unobserved random variables can be written as \( Z_j = \zeta^{-1} (Y_j|x_j, \Theta) \), where \( \Theta = \{\{\gamma, \alpha, \sigma_n^2, \sigma_n^2\}, \{\eta, \sigma_q^2, \sigma_q^2, \sigma_p^2\}, \{\lambda, \mu\}\} \) is denoted as the set of structural parameters.

If we assume that \( Z_j = (\eta_{j1}, \eta_{j2}, q_{js}, q_j)' \sim MVN (U, \Sigma) \), where \( MVN \) stands for multivariate normal, then the pdf of \( Z_j \) is given by

\[
g_z (Z_j) = \frac{1}{2\pi} |\Sigma|^{-\frac{1}{2}} \exp \left( -\frac{1}{2} Z_j' \Sigma^{-1} Z_j \right)
\]

Then the concentrated log-likelihood function is

\[
\hat{L} = \sum_{j=1}^{J} \ln \left| \frac{\partial \zeta^{-1}(Y_j|x_j, \Theta)}{\partial Y_j} \right| - \frac{J}{2} \ln \left| \frac{1}{J} \sum_{j=1}^{J} \zeta^{-1}(Y_j|x_j, \Theta) \zeta^{-1}(Y_j|x_j, \Theta)' \right|
\]

The maximum likelihood estimation problem is formulated as

\[
\max_{\Theta} \hat{L} (Y, x; \Theta, \zeta (\Theta))
\]

and ML estimator is defined as

\[
\Theta^{MLE} = \arg \max_{\Theta} \left\{ \max_{\zeta(\Theta)} \hat{L} (Y, x; \Theta, \zeta (\Theta)) \right\}
\]
5.2 Estimation Method

To evaluate the likelihood function, I have to solve the advertising policy function $a^*_j = A_j(q_{js}, x_j)$ as the equilibrium result of the incomplete information game between studios and consumers. In this paper, I assume movie studios of new movies make their advertising decisions in the context of monopolistic competition game; therefore, they make decisions for each movie independently without taking their rivals’ reaction into account explicitly. However, since the equilibrium advertising strategies for movies with all quality levels impact consumers learning about movies’ quality; therefore, studios need to consider the equilibrium strategies of all movie types to make their own advertising decisions. This requires computing the equilibrium strategies of all movie studios as the fixed points of the best response system, as well as solving each studio’s profit maximization problem given that all studios play the equilibrium strategies.

One option is to use the nested fixed-point (NFXP) algorithm proposed by Rust (1987) to solve the maximum likelihood problem defined in formula (9). The general idea about implementing the NFXP algorithm is that it involves two loops: in the outer-loop, search the structural parameter space over $\Theta$ to maximize $\max_{\zeta(\Theta)} \ln L(Y; x; \Theta, \zeta(\Theta))$; in the inner-loop, for any given values of $\Theta$, solve the optimization problems of all agents and find all possible Bayesian-Nash equilibria. When there is more than one equilibrium existing, I have to evaluate the corresponding likelihood value for each equilibrium and choose the one which yield the highest likelihood value. The whole process continues until the outer loop converges. However, applying NFXP algorithm for my model meets some challenges. First, solving the model can be difficult and even impossible for some guess of parameters $\Theta$, and finding all possible equilibria for any guess of structural parameters can be even more computationally difficult. Second the likelihood function as the objective function of the maximization problem can be potentially discontinuous, since for different guesses of $\Theta$, the number of possible equilibria can be different. And it is very hard to find a reliable and efficient numerical method to solve optimization problems with discontinuous functions.

Another option to estimate games like the one presented in this paper is to use two-step estimators (e.g. Bajari, Benkard and Levin (2007)) which are computationally easier than NFXP. Two-step estimators do not require solving for equilibria and, instead, estimate the equilibrium as the nonparametric functions of data. Therefore, it reduces the cost of computation dramatically. However, the performance of two-step estimators suffers from the small sample bias problem in the first step and do not deal with unobservable variables easily.

In this paper, I apply a new constrained optimization approach proposed by Su and Judd (2012), which is referred to as the mathematical program with equilibrium constraints (MPEC) approach. The constrained optimization approach does not require
repeatedly solving for an equilibrium or all the equilibria at each guess of structural parameters. Instead, equilibrium outcomes can be viewed as constraints that only need to hold at the optimum. The structural parameters and endogenous economic variables are chosen so as to maximize the likelihood of the data subject to the constraints that endogenous economic variables are consistent with an equilibrium for the structural parameters. Thus, this approach reduces the perceived computational burden of implementing the maximum-likelihood estimator. Su and Judd (2012) and Su (2014) provide more details about the constrained optimization approach and the comparison of different approaches discussed here. Therefore, the maximum likelihood estimation problem presented in 9 can be reformulated as a constrained optimization problem in the joint space of structural parameters and economic equilibrium as the following:

$$\max_{\Theta,Z, \zeta(\cdot)} \ln L (Y, x; \Theta, \zeta(\Theta))$$

s.t. \[ Y = \zeta (Z|x, \Theta) \]

where equilibrium equations \( Y = \zeta (Z|x, \Theta) \) are written as constraints, and structural parameters \( \Theta \), unobservable variables \( Z \) and Bayesian-Nash equilibrium \( \zeta (\cdot) \) are chosen to maximize the objective function. The difficulty of the MPEC method (and constrained optimization in general) depends more on convexity and sparsity than the number of unknown parameters. Therefore, instead of solving \( Z = \zeta^{-1} (Y|x, \Theta) \) for each observation and each guess of \( \Theta \), I choose optimal values for \( Z \) which both maximize the objective function and satisfy the equilibrium equation constraints to reduce computational burden.

By solving the optimization problems of both demand and supply sides, I can derive the equilibrium equations \( Y = \zeta (Z|x, \Theta) \) from the model. The number of tickets sold in both periods (equations (5), (6)) and the post-release advertising function (equation (7)) have closed-form expressions, but the pre-release advertising policy function \( a^*_j = A_1 (q_{js}; x_j, \Theta) \) cannot be written in an analytical format explicitly. However, the first-order condition with respect to \( a^*_j \) presented in equation (8) can be used as the equilibrium condition which determines the pre-release advertising policy function of a studio in an environment described by the structural parameter vector \( \Theta \). Determining the exact equilibrium advertising policy function requires solving the advertising first-order condition at an infinite number of values of the private signal received by studios, \( q_{js} \), conditional on each observed value of \( x_j \). To reduce computational burden, instead of considering all possible values of \((q_{js}, x_j)\), the studios’ first-order conditions are solved at a subset of points in the support of \((q_{js}, x_j)\) and the policy function is approximated using Chebyshev polynomials\(^{19}\).

Since consumers are assumed to be able to understand the signaling mechanism and

\(^{19}\)Chebyshev polynomials are used to maximize the stability of the approximation to the policy functions and avoid Runge’s oscillatory phenomenon.
infer \( q_{js} \) from observed advertising intensity and market share. The equilibrium advertising policy function influences consumers’ watching decisions through their utility function. Therefore, instead of approximating advertising policy function, I approximate the inverse function, \( q_{js} = Q(q_{js}, x_j) \), of \( a_j^* = A(q_{js}, x_j) \). Further, all observed characteristics of a new movie included in \( x_j \) enter the consumer’s utility function as a linear combination, so they should enter the advertising policy function in the same way in the equilibrium as well. I define their linear combination as \( X_j = \gamma x_j \), and the inverse advertising policy function becomes a function of two state variables, which reduces the computation challenge dramatically. Let \( na \) be the order of Chebyshev polynomials for \( a_{j1} \) and \( nX \) be the order of Chebyshev polynomials for \( X_j \). The inverse advertising function \( q_{js} = Q(a_{j1}, X_j) : [a_1, \overline{a}_1] \otimes [\overline{X}, \overline{X}] \to R \) is approximated by \( \hat{Q}(a_{j1}, X_j) = \kappa' \lambda(a_{j1}, X_j) \), where \( \kappa \) is \( N \times 1 \) vector of approximation parameters, \( \lambda(a_{j1}, X_j) \) is \( N \times 1 \) vector of \( N \) Chebyshev polynomials and \( N = (na + 1)(nX + 1) \). \( a_m, x_m \) are grids of \( ma \geq na \) and \( mX \geq nX + 1 \) Chebyshev nodes on \([a_1, \overline{a}_1] \otimes [\overline{X}, \overline{X}]\).

With \( q_{js} = \hat{Q}(a_{j1}, X_j) \), the equilibrium equation (8) can be approximated by

\[
\frac{\partial \Pi \left( a_{j1}, \hat{Q}(a_{j1}, X_j); x_j, \Theta \right)}{\partial a_{j1}} + \frac{\partial \Pi \left( a_{j1}, \hat{Q}(a_{j1}, X_j); x_j, \Theta \right)}{\partial \hat{Q}(a_{j1}; X_j)} \frac{\partial \hat{Q}(a_{j1}; X_j)}{\partial a_{j1}} \approx 0 \tag{10}
\]

where \( \Pi \left( a_{j1}, \hat{Q}(a_{j1}, X_j); x_j, \Theta \right) = \pi_{j1} (a_{j1}; SV_{j1}) + E_{j1} V(SV_{j2}|a_{j1}, SV_{j1}) \). Note that the pre-release advertising \( a_{j1} \) has two effects: the first part of the equation shows the “reaching effect” of advertising and the second part of the equation shows the “signaling effect” of advertising.

The maximum likelihood estimation problem formulated as a constrained optimization problem is presented as

\[
\max_{\Theta, \{z_j\}_{j=1}^\kappa} \hat{L}(Y, x; \Theta, \zeta(\Theta))
\]

s.t. \( q_{js} = \hat{Q}(a_{j1}, X_j) = \kappa' \lambda(a_{j1}, X_j) \)

\( a_{j1} = S_1 \left( a_{j1, \eta_{j1}}; x_j, \Theta, \hat{Q}(a_{j1}, X_j) \right) \)

\( a_{j2} = A_2 \left( q_j; a_{j1}, \delta_{j1}, x_j, \Theta, \hat{Q}(a_{j1}, X_j) \right) \)

\( s_{j2} = S_2 \left( a_{j2}, q_j, \eta_{j2}; a_{j1, \eta_{j1}}; x_j, \Theta, \hat{Q}(a_{j1}, X_j) \right) \) for \( j = 1, 2, \ldots, J \)

\[
0 = \frac{\partial \Pi \left( a_{m1}, \hat{Q}(a_{m1}; X_m); X_m, \Theta \right)}{\partial a_{m1}} + \frac{\partial \Pi \left( a_{m1}, \hat{Q}(a_{m1}; X_m); X_m, \Theta \right)}{\partial \hat{Q}(a_{m1}; X_m)} \frac{\partial \hat{Q}(a_{m1}; X_m)}{\partial a_{m1}}
\]

for \( m = 1, 2, \ldots, (ma \times mX) \)
Here, the structural parameters, $\Theta$, unobservables, $\left\{ \eta_j, \eta_j, q_j, s_j \right\}_{j=1}^J$, and approximation parameters, $\kappa$, are chosen to maximize the likelihood function. Integrals over demand shocks in period 1, $\eta_{j1}$, and true unobserved quality, $q_j$, are approximated by using 20 draws from their distributions with antithetic acceleration. Integrals over demand shocks in period 2, $\eta_{j2}$, and WOM signal $q_{ijw}$, are approximated by using 20 draws from their distributions with antithetic acceleration. Integrals over demand shocks in period 2, $\eta_{j2}$, and WOM signal $q_{ijw}$, are approximated by using Gauss–Hermite quadrature with 4 points to improve the speed of estimation. The two-dimensional Chebyshev approximation used for the optimal advertising policy function has 4 degree of Chebyshev polynomials for $a_{j1}$ and 3 degree of Chebyshev polynomials for $X_j$. Then the total number of approximation parameters to be estimated is 20. Equilibrium condition (10) is evaluated at $m_a = 5$ points in the domain of $[0.68, 42]$ for $a_1$ and $m_X = 4$ points in the domain of $[0, 2.5]$ for $X_j$. To increase the computation accuracy and reduce the estimation time, I provided the first-order analytical derivatives of the objective function and constraints and the sparsity pattern of the constraint Jacobian.

5.3 Identification

The dataset provides several sources of variation across movies and weeks to identify the structural parameters $\Theta = \left\{ \{\gamma, \alpha, \sigma_{q1}^2, \sigma_{q2}^2\}, \{\bar{q}, \sigma_{q1}^2, \sigma_{q2}^2, \sigma_{s}^2, \sigma_{d1}^2\}, \{\lambda, \mu\} \right\}$. There are three types of parameters: demand preference parameters $\{\gamma, \alpha, \sigma_{q1}^2, \sigma_{q2}^2\}$; information structure parameters $\{\bar{q}, \sigma_{q1}^2, \sigma_{q2}^2, \sigma_{s}^2, \sigma_{d1}^2\}$; advertising parameters (supply side parameters) $\{\lambda, \mu\}$. I will discuss their identification in turn.

Since only data from one market (the U.S. domestic market) over time can be observed for each movie, preference parameters for observed characteristics, $\gamma$, are assumed to be the same for every consumer or can be taken as the parameters for the average consumer. The variance in advertising spending $a_{j1}$ and $a_{j2}$ corresponding to the variance in $x_j$ can be used to identify $\gamma$. As mentioned before, I divide time into two periods: opening week and post-release weeks to make estimation easier. For post-release weeks, I aggregate the box office performance and advertising spending data together and use a time dummy variable to capture the demand’s level difference between period 1 and 2 due to the difference in time duration. Conditional on $a_{j1}$, $a_{j2}$ and $x_j$, the level difference between $s_{j1}$ and $s_{j2}$ across all movies can be used to identify the coefficient for time dummy, $\alpha$. For aggregate demand shocks, the variance in $s_{j1}$ conditional on $x_j$ and $a_{j1}$, can be used to identify $\sigma_{q1}^2$, and the variance of $s_{j2}$ conditional on $s_{j1}$, $a_{j1}$, $x_j$, and $a_{j2}$ can be used to identify $\sigma_{q2}^2$. The distribution parameters, $\{\bar{q}, \sigma_{q1}^2\}$, for movie’s unobserved quality, $q_j$, can identified by mean and variance of $a_{j2}$ conditional on $x_j$, $a_{j1}$ and $d_{j1}$. The noisy signal

---

20 Quadratures are used instead of simulation because they perform much better when compared with the results of simulation, and allowed for much faster execution.

21 For the domain of $a_{j1}$, I use the observed range of advertising spending in period 1 in the data. For the domain of $X_j$, I first try a large enough range and then adjust the range to appropriate values to improve the estimation accuracy.
variance $\sigma^2_v$ can be identified by the variance of $a_{j1}$ conditional on $a_{j2}$ after $(\bar{q}, \sigma^2_q)$ become known. The adjusted WOM variance parameter $\frac{\sigma^2_w}{\rho}$ can be identified by covariance of $s_{j1}$ and $s_{j2}$ conditional on $a_{j1}$ and $a_{j2}$. Note that the information transmission speed parameter, $\rho$, cannot be separately identified from WOM variance $\sigma^2_w$, but assuming $\sigma^2_w$ as a constant over time is a reasonable assumption. “Reach-efficiency” parameter $\lambda$ in the advertising reach function $\varphi_t(\cdot)$ can be identified by covariance of $s_{j2}$ and $a_{j2}$ conditional on $\tau_2(\cdot)$. Then “advertising depreciation parameter” $\mu$ can be identified by covariance of $s_{j2}$ and $a_{j1}$ conditional on $\tau_2(\cdot)$ when $\lambda$ becomes known.

6 Empirical Results

6.1 Estimates

6.1.1 Advertising’s Signaling Effect

The estimated inverse advertising policy function is presented in Figure 6 and Table 3. For the constrained MLE estimation, no shape constraint is imposed on advertising function, but only the first order conditions of studios’ profit optimization problem are required to be satisfied. Figure 6 shows that, conditional on observed quality ($X$), the unobserved quality is an increasing function of advertising spending $a_1$ for almost all values of $X$, except when $X$’s value is very close to its upper bound. With more details about the inverse advertising policy function, Table 3 shows that only when $a_1$ is very low and $X$ is very high, the “U” shape curve happens (which is highlighted in green). However, only the scenarios shown in the lower right corner of Table 3 are found in the data: movies with high value of $X$ usually have high value of $q_s$ as well.

The estimated advertising policy being an increasing function of $q_s$ conditional on $X$ makes it possible that advertising can play a signaling role if consumers are aware of the positive correlation between $a_1$ and $q_s$. For the model estimated, consumers are assumed to be fully rational so that they understand the signaling mechanism and infer unobservable product quality from the advertising. On the other hand, consumers may be limited rational, which means that they only learn about product existence and observable attributes through advertising. In this case, equation (1) becomes

$$E_{ij1}^c [q_j | I (1)] = \bar{q}$$

and advertising is only used to reach consumers.

The model with limited rationality assumption about consumers is estimated and the corresponding maximized likelihood is compared with the one from the original model. In a comparison of likelihood values, the original model is preferred, which supports the
existence of advertising’s signaling effect.

Utility Function Parameters The estimates of the preference parameter in consumer utility function is reported Table 4.1. Most coefficients of observed characteristics are significant and have the expected signs. In general, movies with higher budgets and higher critic reviews attract consumers more. More movies released widely in the same week makes it tougher for a particular movie to compete for consumers. It seems that consumers get higher utility when watching movies with longer runtime. Movies released by major distributors are much more preferred by consumers in general, compared with those released by mini-major distributors and others. An average consumer obtains more utility from movies with “action” and “comedy” elements and less from movies with “horror” element. Movies rated as “PG” and “PG-13” by MPAA attract more consumers compared with those rated as “G” and “R”. It is not surprising that the coefficient for time dummy is significantly positive, considering the longer period of time in the post-release period. The coefficients for two season indicators, “summer” and “holiday”, are not significant, which seems contradict with the observed strong seasonality of the movie industry in the data. Einav (2007) decomposes the observed seasonal pattern of sales into two components: the underlying demand and seasonal variation in the quality of movies released. He finds that the estimated seasonality in underlying demand is much smaller and slightly different from the observed seasonality of sales after controlling the quality of movies. To some extent, my results are consistent with his arguments.

Information Learning Parameters Table 4.2 presents the estimated parameters for information learning about a new movie’s quality. The prior distribution for \( q \) has a mean equaling to -0.317 and variance equaling to 0.339. The interpretation of the prior distribution variance is that, before any information available to either a studio or consumers to learn about a new movie’s true unobserved quality, both parties face an uncertainty (measured by the relative standard deviation) of 187.8% of the systematic quality. The variance of the noisy signal, \( \sigma^2 \), measures how accurately studios can learn about a new movie’s quality through marketing research before releasing it. The higher the value, the less efficient their marketing research is. The estimated \( \sigma^2 \) equaling 3.501 means that studios’ marketing research doesn’t help them learn much about the movie’s quality. When a studio updates its belief about a new movie’s quality, the weight it should put on the received noisy signal is \( \beta^q (0) = \frac{\sigma^2}{\sigma^2(0)+\sigma^2} = \frac{\sigma^2}{\sigma^2+\sigma^2} = 0.09 \), and the updated variance becomes \( \sigma^2 (1) = \frac{\sigma^2+\sigma^2}{\sigma^2+\sigma^2} = 0.309 \). On the other hand, the adjusted variance of WOM is only 0.023. This indicates that WOM among consumers is much more efficient and dominant communication channel to reveal information about a movie’s true quality. On

\[^{22}\text{Here I measure the relative quality of models for a given set of data by using Akaike Information Criterion (AIC).}\]
average\(^\text{23}\), consumers put around \(\beta_q^e(1) = \frac{\sigma_q^2(1)}{\sigma_q^2(1) + \sigma_\delta^2} = 0.93\) weights on WOM information and only 0.07 weights on firm generated information after the release of a new movie and the updated variance is \(\sigma_q^2(2) = \frac{1}{\sigma_q^2(1) + \sigma_\delta^2} = \frac{1}{0.93^2 + 0.07^2} = 0.02\) in the post-release weeks which is much smaller than prior variance.

### Advertising Reaching Function Parameters

Parameters for the advertising reach function are presented in Table 4.3. With \(\lambda\) equaling to 0.131, the market coverage ranges from 8.55% to 99.6% for movies in our dataset. With the average 15 million dollars advertising spending in period 1, around 86% of the market is covered. This shows the advertising in this industry has high reaching efficiency. However, the value of \(\mu\) shows that it is easy for consumers to forget about the movie. The depreciation rate for the advertising stock \((1 - \mu)\) is 0.684, so only 31.6% of advertising spending in period 1 still works in period 2. By that time, a great proportion of advertising spending is actually used to remind consumers about the movie, instead of reaching for new consumers.

### 6.2 Model Fit

To examine the robustness of the estimated model, I conduct several goodness-of-fit tests to check how well the predicted data generated by the model fits the observed data from my sample. More specifically, I am interested in how well the model predicts studios’ advertising choices (both pre-release and post-release advertising expenses) and how well it predicts consumers’ choices (the number of tickets sold both in the opening week and post-release weeks).

Based on the estimated parameters of the structural model, I simulate a large number of advertising spending and box office performance over time for each movie in my sample. Then I partition the region in which each interested response variable lies into 5 disjoint cells. By construction, the observed values of the interested response variable have 20% probability to fall into each cell. In general, the test statistic is of the form:

\[
X^2 = \sum_{k=1}^{K} \frac{(n_k - n_k^e)^2}{n_k^e} \sim X^2_{K-1}
\]

where \(n_k\) is the number of observations that fall into cell \(k\), and \(n_k^e\) is the number of observations that the model predicts should be in cell \(k\). \(n_k^e = p_kN\) where \(N\) is the observed sample size and \(p_k\) is calculated by using the simulated data. The test statistic approximately follows a chi-square distribution with \(K - 1\) degree of freedom where \(K\) is the number of cells.

Table 5-1 shows the observed and expected numbers of data fall in each cells for both

\(^{23}\text{Mean}(\delta_1)\) is normalized to 1.
advertising spending and box office performance in two periods. The null hypothesis of the formal test is that there is no difference between the observed advertising spending (box office performance) and the predicted advertising spending (box office performance). The 10% level of significance critical value of the chi-square distribution with 4 degree of freedom is 7.78. Therefore, in general, the model fits data well.

To further examine how well the estimated pre-release advertising policy function performs, I check how well the model predict studios’ pre-release advertising spending conditional on different values of observed attributes. I nonparametrically partition \( a_1 \) and \( X = x \) separately and form cross-product cells, where \( a_1 \) is the pre-release advertising spending and \( X = \gamma x \) is the linear combination of observable attributes of a movie. Then I calculate the chi-square statistic conditional on each cell of \( X \). Table 5-2 displays the fit of pre-release advertising conditional different value ranges of \( X \). Controlling the movie’s observable attributes, the model does a good job of predicting pre-release advertising spending across cells. However, the model tends to fit the data less well when \( X \)’s value is high.

7 Counterfactual Analysis

The goal of the counterfactual experiments in this section is to understand how studios’ advertising spending decisions are affected by consumer information learning through different channels. Specifically, I try to 1) separate advertising’s signaling effect from its reaching effect to understand how much lower advertising spending would be if advertising was only used to reach consumers, and 2) understand how studios’ advertising spending allocation over time would be under different information structures.

The setup in equations (5) and (6) shows that advertising affects demand through two channels: how extensively advertisements reach consumers and how consumers take advertising intensity as quality signals. Ideally, we can consider a world where consumers automatically have the same information about a new movie’s quality as consumers in the estimated model, without inferring from advertisement intensity (in period 1) or market performance (in period 2). In that case, studios only use advertising to reach consumers, not to signal movie quality. However, in the estimated model, consumers only observe advertising intensity in period 1, and that brings some noisiness to consumer learning and also makes the simulation exercise difficult. Alternatively, we can consider a world where consumers who are reached by advertisements automatically know \( q_s \) as well as studios when making a decision and do not need to infer any information about \( q_s \) from advertisement intensity (in period 1) or market performance (in period 2). Likewise, we can consider a world where consumers who are reached by advertisements do not know \( q_s \) when making a decision and have limited rationality towards information learning through advertisement intensity or market performance. In both cases, there is no learn-
ing from studios’ actions, eliminating the need for signaling effect of advertising. The fact that consumers are either perfectly informed or uninformed makes the ideal case fall somewhere in between these two cases. The differences in advertising strategies and spending between these two cases and the actual advertising strategies and spending, give us an idea about the amount of advertising money that is spent for signaling and reaching purposes separately as well as how studios’ optimal advertising strategies are affected by the information structure of this industry.

**Experiment 1 (Exp1):** No information asymmetry about \( q_s \) and therefore no advertising signaling needed.

In the estimated model, only the studio receives \( q_{js} \) before the release of its new movie \( j \), however, after the release of the movie, consumers fully accept and analyze all available information implied by the studio’s actions and infer the movie’s quality through them. Therefore, consumer \( i \)’s perceived expected quality of movie \( j \) is

\[
E_{ij1}^c[q_j | k_{ij1}, A_1(q_{js})] = \int q_j \cdot g_{i1}(q_j|k_{ij1}, A_1(q_{js})) \, dq_j
\]

for period 1 and

\[
E_{ij2}^c(q_j) = E_{ij1}^c[q_j | d_{j1}] + \beta_q^c(1) \left( \bar{q}_{ijw} - E_{ij1}^c[q_j | d_{j1}] \right)
\]

for period 2. Both of them are affected by advertising spending \( a_{j1} \) through its signaling effect. The demand in period 1, \( d_{j1} = \varphi_1(a_{j1}) \tau_1(a_{j1}; \cdot) M \), shows that \( a_{j1} \) affects demand both through reaching channel (\( \varphi_1(a_{j1}) \)) and signaling channel (\( \tau_1(a_{j1}; \cdot) \)).

For counterfactual experiment 1, I assume consumers automatically know \( q_{js} \) as well as the studio after the movie is released. Therefore, consumer \( i \)’s perceived expected quality of movie \( j \) becomes

\[
E_{ij1}^c(q_j) = E_{i1}^c(q_j) = \bar{q} + \beta_q^c(0) \left( q_{js} - \bar{q} \right)
\]

in period 1 and

\[
E_{i2}^c(q_j) = E_{i1}^c(q_j) + \beta_q^c(1) \left( \bar{q}_{ijw} - E_{ij1}^c(q_j) \right)
\]

in period 2. Both are independent of studios’ advertising spending and the equilibrium advertising strategies. The demand in period 1 becomes \( d_{j1} = \varphi_1(a_{j1}) \tau_1(\cdot) M \) which \( a_{j1} \) affects only through reaching channel (\( \varphi_1(a_{j1}) \)).

**Experiment 2 (Exp2):** There is information asymmetry about \( q_s \) but consumers are limited rational towards information learning.

For this experiment, I assume that consumers don’t know \( q_{js} \) while the studio knows it, and consumers do not infer the movie’s quality information from the received advertising intensity or the market performance. Therefore, consumer \( i \)’s perceived expected quality
of movie $j$ becomes

$$E_{ij1}^c(q_j) = \bar{q}$$

in period 1 and

$$E_{ij2}^c(q_i) = \bar{q} + \beta_q^c(1)(\bar{q}_{ijw} - \bar{q})$$

in period 2. Similar to experiment 1, $a_{j1}$ affects the demand only through its reaching effect.

The results of these two experiments are reported in Table 6 and Figure 7. For all 632 movies in my sample, the total advertising spending for both pre-release and post-release stages is around $13$ billion. For both simulated cases, when advertising is only used to reach consumers, the total advertising spending is around $9.5$ billion, only 73% of the original case. Therefore, after teasing out the reaching effect of advertising, we see that around 27% of all the advertising money for movies in my sample is “burned” for the signaling purpose. If we examine how studios allocate advertising money over time, it is very different for the original case and for the simulated cases. When advertising plays both signaling and reaching roles, on average, about 76% of the total advertising budget is spent in the pre-release stage. When advertising is only used for reaching consumers, on average, advertising money is arranged much more evenly over time, with roughly 50% spent in the pre-release stage and another 50% spent in the post-release stage.

In Table 6, the advertising spending pattern is very similar for experiment 1 and experiment 2. This is because only the average advertising spending over all movies in my sample is shown in Table 6. For movies with different characteristics (observed and unobserved), studios’ advertising strategies are very different under those two different information structures. In Figure 7, I simulate advertising strategies for all three cases with different values of $q_s$ and $X$. Conditional on $X$, movies with high $q_s$ have higher pre-release advertising spending and lower ex-ante expected post-release advertising spending in experiment 1 than in experiment 2. Movies with low $q_s$ have lower pre-release advertising spending and higher ex-ante expected post-release advertising spending in experiment 1 than in experiment 2. Intuitively, when consumers have full information about $q_s$, studios with high $q_s$ movies would like to spend more in the pre-release stage, since consumers understand they have high $q_s$ and therefore more consumers tend to watch the movie. When consumers have no information about $q_s$, studios with low $q_s$ movies want to spend more in pre-release stage, since consumers can not differentiate their movie from movies with high $q_s$ in opening week. In this way, studios with low $q_s$ movies can recoup as much of their investment as possible before consumers realize the low quality of their movies later after learning this through WOM. When the $X$ value is low, full information about $q_s$ even prevents movies with very low $q_s$ from entering the market. If we take experiment 1 as the “full information” case and experiment 2 as the “no information” case, then the estimated case can be taken as a “signaling” case, which reduces the information
asymmetry between studios and consumers through advertising signals. Figure 8 gives an example of the comparison of those three cases.

8 Conclusion and Future Work

For experience goods where information asymmetry exists between firms and consumers, the interaction between advertising and WOM communication among consumers have not been fully explored. This paper models movie studios’ optimal advertising strategies and consumer information learning in an equilibrium setting in the motion picture industry of the United States. In my model, advertising conveys direct information about a movie’s existence and attributes in addition to signaling its quality. The WOM mechanism is used to constrain studios’ behavior and fulfills the signaling role of advertising. I use weekly data from the U.S. movie theatrical market to empirically test and measure the signaling effect of advertising. By distinguishing between two types of informative advertising effects, I find that around 27% of advertising spending on the movies in my sample is “burned” for a signaling purpose, while 73% of advertising money is spent to reach consumers.

Studios’ advertising strategies over time differ when advertising is used only to reach consumers, with around 50% spent in the pre-release stage. When studios need to use advertising to signal movie quality, they allocate 76% of money for pre-release advertising. I also try to quantify how much value the “money-burning” advertising can produce, in terms of reducing information uncertainty faced by consumers, by scrutinizing movies of varying quality levels.

The estimated information parameters (prior-and-post variances of expected movie quality) from my model also show that studios usually fail to learn effectively about their movies’ true quality, while WOM reveals the true quality of a movie to consumers more efficiently. In the post-release weeks, the uncertainty about a movie’s quality is greatly reduced by more than 90%, mainly through the WOM channel. By conducting a set of counterfactual experiments, I evaluate the value of information learning for studios through pre-release marketing research and for consumers through post-release WOM.

In this paper, I use a simplified two-period model to capture the change of information structure before and after a movie is released. One possible extension is to set up a multiple-period model by weeks, which may capture more features of firms’ dynamic optimal decisions. Another logic extension of this study is to consider the impact of revenue from nontheatrical windows on studios’ optimal advertising decisions for theatrical window. In this paper, I assume that studios aim to run the U.S. theatrical release window in a stand-alone profitable manner. However, nontheatrical windows, especially the home video window, have emerged as very profitable ones, and studios may consider the theatrical window as an advertisement for the nontheatrical windows. Therefore, the al-
ternative assumption is the studios optimize advertising spending across multiple release windows. Besides, I propose a new method to model how advertising reaches consumers and simultaneously signals product quality, which can be generalized to other industries.

References


### Table 1: Summary Statistics of Main Variables

<table>
<thead>
<tr>
<th>Category</th>
<th>Variables</th>
<th>Mean</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-release ads</td>
<td>ad1</td>
<td>15.698</td>
<td>6.608</td>
</tr>
<tr>
<td></td>
<td>ad2</td>
<td>4.965</td>
<td>4.0878</td>
</tr>
<tr>
<td>Total ads</td>
<td>adt</td>
<td>20.663</td>
<td>9.676</td>
</tr>
<tr>
<td>Opening week BOR</td>
<td>bor1</td>
<td>14.128</td>
<td>12.198</td>
</tr>
<tr>
<td>Post-release weeks BOR</td>
<td>bor2</td>
<td>33.400</td>
<td>35.451</td>
</tr>
<tr>
<td>Total BOR</td>
<td>bort</td>
<td>47.528</td>
<td>46.353</td>
</tr>
<tr>
<td>Production Budget</td>
<td>budget</td>
<td>44.593</td>
<td>31.907</td>
</tr>
<tr>
<td>Season</td>
<td>holiday</td>
<td>0.109</td>
<td>0.312</td>
</tr>
<tr>
<td></td>
<td>summer</td>
<td>0.285</td>
<td>0.452</td>
</tr>
<tr>
<td>MAPP</td>
<td>G</td>
<td>0.030</td>
<td>0.171</td>
</tr>
<tr>
<td>Ratings</td>
<td>PG</td>
<td>0.153</td>
<td>0.360</td>
</tr>
<tr>
<td></td>
<td>PG-13</td>
<td>0.486</td>
<td>0.500</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>0.330</td>
<td>0.470</td>
</tr>
<tr>
<td>Genres</td>
<td>action</td>
<td>0.142</td>
<td>0.349</td>
</tr>
<tr>
<td></td>
<td>comedy</td>
<td>0.397</td>
<td>0.490</td>
</tr>
<tr>
<td></td>
<td>drama</td>
<td>0.162</td>
<td>0.369</td>
</tr>
<tr>
<td></td>
<td>family</td>
<td>0.132</td>
<td>0.339</td>
</tr>
<tr>
<td></td>
<td>horror</td>
<td>0.106</td>
<td>0.308</td>
</tr>
<tr>
<td>Distributors</td>
<td>major</td>
<td>0.765</td>
<td>0.424</td>
</tr>
<tr>
<td></td>
<td>Mini-major</td>
<td>0.128</td>
<td>0.334</td>
</tr>
<tr>
<td></td>
<td>Others</td>
<td>0.107</td>
<td>0.310</td>
</tr>
<tr>
<td>Critic</td>
<td>metacritic</td>
<td>45.009</td>
<td>16.662</td>
</tr>
<tr>
<td># of Competitors</td>
<td>ncompete</td>
<td>2.320</td>
<td>1.162</td>
</tr>
<tr>
<td>Ticket Price</td>
<td>price</td>
<td>5.915</td>
<td>0.344</td>
</tr>
<tr>
<td>Runtime (minutes)</td>
<td>runtime</td>
<td>105.074</td>
<td>16.570</td>
</tr>
</tbody>
</table>

Note: this table uses the sample of 632 movies released between Feb., 2000 and Nov., 2005; Advertising spending, box office revenue and production budget are all in millions.
Table 2: Short-run Vs Long-run Market

<table>
<thead>
<tr>
<th></th>
<th># of obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>All movies in the sample:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Market share for opening week</td>
<td>632</td>
<td>0.82%</td>
<td>0.70%</td>
<td>0.04%</td>
<td>4.18%</td>
</tr>
<tr>
<td>Market share for post-release</td>
<td>632</td>
<td>1.95%</td>
<td>2.08%</td>
<td>0.02%</td>
<td>15.39%</td>
</tr>
<tr>
<td>Pre-release ads/Total ads&gt;95%:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total BOR/ Total ads spending</td>
<td>20</td>
<td>63.29%</td>
<td>30.77%</td>
<td>17.50%</td>
<td>126.46%</td>
</tr>
<tr>
<td>Total ads spending</td>
<td>20</td>
<td>11.01</td>
<td>4.06</td>
<td>2.77</td>
<td>19.31</td>
</tr>
<tr>
<td>Total BOR</td>
<td>20</td>
<td>6.85</td>
<td>3.77</td>
<td>1.07</td>
<td>14.38</td>
</tr>
<tr>
<td>Pre-release ads/Total ads</td>
<td>20</td>
<td>97.02%</td>
<td>1.54%</td>
<td>95.02%</td>
<td>99.84%</td>
</tr>
<tr>
<td>Opening week BOR/Total BOR</td>
<td>20</td>
<td>48.30%</td>
<td>7.94%</td>
<td>32.59%</td>
<td>63.22%</td>
</tr>
<tr>
<td>Pre-release ads/Total ads &lt;60%:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total BOR/ Total ads spending</td>
<td>34</td>
<td>389.67%</td>
<td>193.08%</td>
<td>49.70%</td>
<td>964.69%</td>
</tr>
<tr>
<td>Total ads spending</td>
<td>34</td>
<td>30.51</td>
<td>13.42</td>
<td>7.34</td>
<td>55.90</td>
</tr>
<tr>
<td>Total BOR</td>
<td>34</td>
<td>122.13</td>
<td>84.00</td>
<td>13.54</td>
<td>339.72</td>
</tr>
<tr>
<td>Pre-release ads/Total ads</td>
<td>34</td>
<td>53.17%</td>
<td>5.44%</td>
<td>35.30%</td>
<td>59.98%</td>
</tr>
<tr>
<td>Opening week BOR/Total BOR</td>
<td>34</td>
<td>23.08%</td>
<td>6.97%</td>
<td>12.20%</td>
<td>43.04%</td>
</tr>
</tbody>
</table>

Table 3: Approximated Inversed Advertising Policy Function

<table>
<thead>
<tr>
<th>a1</th>
<th>x</th>
<th>0</th>
<th>0.20</th>
<th>0.40</th>
<th>0.60</th>
<th>0.80</th>
<th>1.00</th>
<th>1.20</th>
<th>1.40</th>
<th>1.60</th>
<th>1.80</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.68</td>
<td>-2.67</td>
<td>-3.61</td>
<td>-4.37</td>
<td>-4.99</td>
<td>-5.50</td>
<td>-5.93</td>
<td>-6.32</td>
<td>-6.70</td>
<td>-7.10</td>
<td>-7.55</td>
<td></td>
</tr>
<tr>
<td>4.96</td>
<td>1.03</td>
<td>-0.35</td>
<td>-1.60</td>
<td>-2.73</td>
<td>-3.77</td>
<td>-4.73</td>
<td>-5.66</td>
<td>-6.55</td>
<td>-7.45</td>
<td>-8.37</td>
<td></td>
</tr>
<tr>
<td>9.24</td>
<td>4.19</td>
<td>2.52</td>
<td>0.96</td>
<td>-0.50</td>
<td>-1.89</td>
<td>-3.21</td>
<td>-4.49</td>
<td>-5.74</td>
<td>-6.97</td>
<td>-8.21</td>
<td></td>
</tr>
<tr>
<td>13.52</td>
<td>6.90</td>
<td>5.06</td>
<td>3.32</td>
<td>1.65</td>
<td>0.04</td>
<td>-1.51</td>
<td>-3.02</td>
<td>-4.49</td>
<td>-5.95</td>
<td>-7.40</td>
<td></td>
</tr>
<tr>
<td>17.80</td>
<td>9.24</td>
<td>7.32</td>
<td>5.46</td>
<td>3.67</td>
<td>1.94</td>
<td>0.25</td>
<td>-1.40</td>
<td>-3.02</td>
<td>-4.62</td>
<td>-6.21</td>
<td></td>
</tr>
<tr>
<td>22.08</td>
<td>11.30</td>
<td>9.32</td>
<td>7.41</td>
<td>5.55</td>
<td>3.74</td>
<td>1.97</td>
<td>0.23</td>
<td>-1.48</td>
<td>-3.17</td>
<td>-4.85</td>
<td></td>
</tr>
<tr>
<td>26.35</td>
<td>13.12</td>
<td>9.17</td>
<td>7.27</td>
<td>5.41</td>
<td>3.59</td>
<td>1.79</td>
<td>0.02</td>
<td>-1.73</td>
<td>-3.47</td>
<td>-5.47</td>
<td></td>
</tr>
<tr>
<td>30.63</td>
<td>14.76</td>
<td>12.74</td>
<td>10.77</td>
<td>8.83</td>
<td>6.93</td>
<td>5.07</td>
<td>3.24</td>
<td>1.42</td>
<td>-0.38</td>
<td>-2.17</td>
<td></td>
</tr>
<tr>
<td>34.91</td>
<td>16.24</td>
<td>14.21</td>
<td>12.21</td>
<td>10.25</td>
<td>8.32</td>
<td>6.42</td>
<td>4.54</td>
<td>2.68</td>
<td>0.84</td>
<td>-0.98</td>
<td></td>
</tr>
<tr>
<td>39.19</td>
<td>17.60</td>
<td>15.54</td>
<td>13.52</td>
<td>11.53</td>
<td>9.57</td>
<td>7.64</td>
<td>5.73</td>
<td>3.84</td>
<td>1.96</td>
<td>0.10</td>
<td></td>
</tr>
</tbody>
</table>
### Table 4.1: Estimated Parameters for Utility Function

<table>
<thead>
<tr>
<th>Coefficients for Observed Characteristics ($\gamma$)</th>
<th>Estimates</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Budget</td>
<td>0.064***</td>
<td>0.0002</td>
</tr>
<tr>
<td>Critic</td>
<td>0.058***</td>
<td>0.0006</td>
</tr>
<tr>
<td>Summer</td>
<td>-0.292</td>
<td>0.3211</td>
</tr>
<tr>
<td>Holiday</td>
<td>0.826</td>
<td>1.0756</td>
</tr>
<tr>
<td># of competitors</td>
<td>-0.030***</td>
<td>0.0016</td>
</tr>
<tr>
<td>Runtime</td>
<td>0.221***</td>
<td>0.0508</td>
</tr>
<tr>
<td>Major</td>
<td>0.129***</td>
<td>0.0123</td>
</tr>
<tr>
<td>Mini-major</td>
<td>0.071***</td>
<td>0.0134</td>
</tr>
<tr>
<td>Action</td>
<td>0.069***</td>
<td>0.0068</td>
</tr>
<tr>
<td>Comedy</td>
<td>1.075*</td>
<td>0.6172</td>
</tr>
<tr>
<td>Drama</td>
<td>-0.125</td>
<td>1.0835</td>
</tr>
<tr>
<td>Family</td>
<td>-0.038</td>
<td>4.8666</td>
</tr>
<tr>
<td>Horror</td>
<td>-0.052***</td>
<td>0.0099</td>
</tr>
<tr>
<td>MPAA_G</td>
<td>0.076</td>
<td>0.1289</td>
</tr>
<tr>
<td>MPAA_PG</td>
<td>0.149***</td>
<td>0.0395</td>
</tr>
<tr>
<td>MPAA_PG13</td>
<td>0.092***</td>
<td>0.0037</td>
</tr>
<tr>
<td>Time dummy ($\alpha$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demand shock variance in period 1 ($\sigma_{\eta_1}^2$)</td>
<td>0.351***</td>
<td>0.0022</td>
</tr>
<tr>
<td>Demand shock variance in period 2 ($\sigma_{\eta_2}^2$)</td>
<td>0.337***</td>
<td>0.0029</td>
</tr>
</tbody>
</table>

Note: 1) *** p<0.01, ** p<0.05, * p<0.1; 2) the sample used includes 632 movies released between Feb., 2000 and Nov., 2005;

### Table 4.2: Estimated Parameters for Information Learning

<table>
<thead>
<tr>
<th>Estimates</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean of quality ($\bar{q}$)</td>
<td>-0.317***</td>
</tr>
<tr>
<td>Variance of quality ($\sigma_q^2$)</td>
<td>0.339***</td>
</tr>
<tr>
<td>Variance of noisy signal of quality ($\sigma_s^2$)</td>
<td>3.501**</td>
</tr>
<tr>
<td>Word of Mouth variance (adjusted) ($\sigma_{wh}^2$)</td>
<td>0.023**</td>
</tr>
</tbody>
</table>

### Table 4.3: Estimated Parameters for Advertising Reach Function

<table>
<thead>
<tr>
<th>Estimates</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reach efficiency parameter ($\lambda$)</td>
<td>0.131***</td>
</tr>
<tr>
<td>Advertising depreciation parameter ($\mu$)</td>
<td>0.316***</td>
</tr>
</tbody>
</table>

Note: 1) *** p<0.01, ** p<0.05, * p<0.1; 2) the sample used includes 632 movies released between Feb., 2000 and Nov., 2005;
### Table 5-1: Data Vs Model

<table>
<thead>
<tr>
<th>Cells</th>
<th>Data</th>
<th>Pre-release Advertising</th>
<th>Post-release Advertising</th>
<th>Opening Week Performance</th>
<th>Post-release Weeks Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cell 1</td>
<td>126</td>
<td>132.61</td>
<td>124.69</td>
<td>139.41</td>
<td>141.83</td>
</tr>
<tr>
<td>Cell 2</td>
<td>127</td>
<td>147.65</td>
<td>110.49</td>
<td>135.26</td>
<td>134.05</td>
</tr>
<tr>
<td>Cell 3</td>
<td>126</td>
<td>122.61</td>
<td>127.03</td>
<td>122.98</td>
<td>134.51</td>
</tr>
<tr>
<td>Cell 4</td>
<td>127</td>
<td>109.83</td>
<td>141.55</td>
<td>128.94</td>
<td>112.29</td>
</tr>
<tr>
<td>Cell 5</td>
<td>126</td>
<td>119.29</td>
<td>128.24</td>
<td>105.41</td>
<td>109.32</td>
</tr>
<tr>
<td>X² Stat.</td>
<td></td>
<td>6.37</td>
<td>4.02</td>
<td>5.92</td>
<td>7.15</td>
</tr>
</tbody>
</table>

Note: cells for different response variables have different ranges;

\[ X²_{0.05} = 7.78 \]

\[ X²_{0.05} = 9.49 \]

### Table 5-2: Data Vs Model: Pre-release Advertising Conditional on Observed Attributes

<table>
<thead>
<tr>
<th></th>
<th>&lt;0.70</th>
<th>[0.70,0.85]</th>
<th>[0.85,0.97]</th>
<th>[0.97,1.15]</th>
<th>&gt;1.15</th>
</tr>
</thead>
<tbody>
<tr>
<td>x</td>
<td>Data</td>
<td>Model</td>
<td>Data</td>
<td>Model</td>
<td>Data</td>
</tr>
<tr>
<td>&lt;9.77</td>
<td>78</td>
<td>70.00</td>
<td>29</td>
<td>35.68</td>
<td>14</td>
</tr>
<tr>
<td>[9.77,13.77]</td>
<td>39</td>
<td>39.88</td>
<td>39</td>
<td>46.16</td>
<td>33</td>
</tr>
<tr>
<td>[13.77,17.18]</td>
<td>8</td>
<td>12.83</td>
<td>38</td>
<td>29.60</td>
<td>42</td>
</tr>
<tr>
<td>&gt;21.44</td>
<td>0</td>
<td>0.16</td>
<td>4</td>
<td>2.33</td>
<td>8</td>
</tr>
<tr>
<td>X² stat</td>
<td>4.36</td>
<td>7.01</td>
<td>2.12</td>
<td>5.42</td>
<td>8.53</td>
</tr>
</tbody>
</table>

\[ X²_{0.05} = 7.78 \]

\[ X²_{0.05} = 9.49 \]

### Table 6: Signaling VS Reaching Role of Advertising

<table>
<thead>
<tr>
<th></th>
<th>Original structure(OS)</th>
<th>Full info.</th>
<th>Full info./OS</th>
<th>No info.</th>
<th>No info./OS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-release ads: a1</td>
<td>9.94</td>
<td>4.67</td>
<td>47%</td>
<td>4.71</td>
<td>47%</td>
</tr>
<tr>
<td>Post-release ads: a2</td>
<td>3.15</td>
<td>4.84</td>
<td>154%</td>
<td>4.81</td>
<td>153%</td>
</tr>
<tr>
<td>Total ads: a1+a2</td>
<td>13.09</td>
<td>9.51</td>
<td>73%</td>
<td>9.52</td>
<td>73%</td>
</tr>
<tr>
<td>a1/ (a1+a2) (%)</td>
<td>75.94%</td>
<td>49.11%</td>
<td>na</td>
<td>na</td>
<td>49.40%</td>
</tr>
<tr>
<td># of movies entering the market</td>
<td>632</td>
<td>630</td>
<td>99.68%</td>
<td>632</td>
<td>100%</td>
</tr>
</tbody>
</table>

Note: All advertising spending is in billions, and all numbers are calculated based on the target sample.
Figure 1: Pre-release and Post-release Advertising Spending VS Production Budget

Figure 2: Sales trends diverge over time after release

Bruno: user rating is 6.1 and box office revenue is $60 million;
District 9: user rating is 8.4 and box office revenue is $116 million.
Figure 3: Game Timing

Period 0

Supply

Up-front assessment → Firm: less uncertain

A new product is produced: quality?

Demand

Period 1

Advertising

Firm: less uncertain

Consumers enter and learn: less uncertain → WOM

Period 2

Advertising

Consumers enter and learn: less uncertain
Figure 4: The Importance of Long-run Market

Note: $K_1^*$ is the threshold level of $K_{ij1}$. When consumers receive more than $K_1^*$ ads, they are convinced to watch the movie.

Figure 5: Distribution of $K_{ijt}$

Note: $K_1^*$ is the threshold level of $K_{ij1}$. When consumers receive more than $K_1^*$ ads, they are convinced to watch the movie.
Figure 6: Inverse Pre-release Advertising Policy Function
Figure 7: Advertising Strategy under Different Information Structure
Figure 8: Comparison of Three Cases (Conditional on X)

- Pre-release advertising spending
- Ex-ante Expected total advertising spending
- Ex-ante Expected Demand
- Ex-ante Expected Profit

Exp1
Exp2
Estimated