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

Haiyan Liu
Sichuan, China

M.A. University of Virginia, 2008
M.A. Peking University, Beijing, China, 2006
B.A. Renmin University of China, Beijing, China, 2003

A Dissertation presented to the Graduate Faculty
of the University of Virginia in Candidacy for the Degree of
Doctor of Philosophy

Department of Economics

University of Virginia
August, 2014

Abstract

When new products are introduced to the market, normally, there exists information asymmetry between consumers and firms. Being uncertain about new products' attributes, consumers are motivated to learn from both from consumer-generated information (e.g. the Word-of-Mouth (WOM) communication among consumers) and firm-generated information (e.g. advertising). When deciding optimal advertising strategies, firms have to be aware how advertising interacts with consumer learning to create an asymmetry in the returns for products with different quality levels.

This dissertation 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. Studios need to decide the pre-release advertising and post-release advertising spending, knowing that the pre-release advertising plays a dual role of informing consumers about a new movie as well as signaling the movie's unobservable quality. Consumers who are reached by advertising and then enter the market at different time have different information sources. Those who enter the market in the opening week, use advertising as the main information source to infer a movie's quality. While those who enter the market in post-release weeks enjoy the extra benefit of information from WOM.

I estimate the model using weekly data on advertising spending, box office performance 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 as well as reaching consumers 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.

Since advertising can be used to reach consumers as well as to signal product quality, I use counterfactual experiments to distinguish the amount of money that is used for purposes of signaling as well as reaching. Counterfactual experiments suggest that around 27% of advertising spending on the movies in my sample is for a signaling purpose, while 73% of advertising money is spent to reach consumers. When signaling movie quality through advertising, studios with high-quality movies tend to spend more and studios with low-quality movies tend to spend less in pre-release weeks than the case when advertising is only used to reach consumers. Information revealed by both advertising signaling and social learning 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 to be "burned" for signaling purposes.

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

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

Acknowledgement

I would like to express my deepest appreciation to my advisors at University of Virginia, Steven Stern, Simon Anderson and Natasha Zhang Foutz. Without their full support and guidance to each stage of my research, I could not have written this dissertation and get through some of the most difficult times of my life over the past few years. My advisors not only taught me the profession I love, their dedication to research, teaching and advising also serve as inspirations for me.

I am also very grateful to Federico Ciliberto, Nathan Larson, David Mills, Michelle Sovinsky Goeree, Ginger Jin and John Rust, as well as many participants of the IO research meeting of Economics Department at University of Virginia, for providing me particularly valuable comments and suggestions which greatly improved this research. I acknowledge the Bankard Fund for Political Economy for financial support and TNS Media Intelligence for generously providing me with the data used in this research.

I want to thank many friends I met at UVA graduate program in the past few years, including Christopher Clapp, Shushanik Hakobyan, Raju Huidrom, Kang Jian, Xiaohuan Lan, Michael Schreck, Ruwei Wang, Shuna Wang, and Yiyi Zhou. Without their professional and personal support and companion, my journey at UVA could be much tougher and less fun.

Most importantly, none of this would have been possible without the love and patience of my family. This dissertation is dedicated to my parents who have supported me to pursue whatever I dream of since I was little, even following my dream sometimes means being far away from them when they need me most. I want to thank my husband and the love of my life, Jianwei, for standing by me through this long and challenging journey. Words cannot describe how much unconditional love, support, encouragement and advices I got from him. To my darling baby, Alex: thank you for

being such a good little baby in the past one year, so that I can finish what I started. Your arrival changed my life and filled me with courage to face any challenges in my life.

Contents

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 that produce 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 dissertation,

[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.

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 strategies 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 dissertation, the data used for estimation comes from widely released movies from 2002 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

[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.

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 dissertation, 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 dissertation, 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 in my model.

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 to 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. To estimate the structural parameters of my model, I write unobserved random variables, such as movie quality and demand shocks, as functions of observed variables such as box office revenue and advertising spending etc.. By assuming that the unobserved variables follow Multivariate Normal distribution, I can get the likelihood function. 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 remainder of this dissertation is organized as follows. In Chapter 2, I begin with a survey of the literature most closely related to my work. In Chapter 3, I briefly describe the U.S. movie theater market and the data used in estimation. Some preliminary results from nonstructural analysis are also presented to get some insight about the signaling effect of advertising in the pre-release stage and the influence of social learning in the post-release stage. Chapter 4 lays out the model. I also define the pure strategy Nash signaling equilibrium and provide some simple examples to discuss the existence of the signaling equilibrium. Chapter 5 explains my empirical strategy, followed by a brief discussion of identification. Chapter 6 presents the estimation results. Chapter 7 conducts the counterfactual experiments. Chapter 8 concludes the paper with some discussion about future work.

2 Literature Review

This dissertation mainly focuses on how advertising interacts with WOM and provides information to help solving information asymmetry problems between firms and consumers. Therefore, in this chapter, I will review the most relevant theoretical and empirical literature of informative advertising first. Then I will briefly describe several empirical studies which are related to advertising, social learning and the motion picture industry.

2.1 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 group of papers focuses on how advertising conveys "hard" (direct) information about a product's existence and attributes. In those studies, consumers may have imperfect information about the availability and attributes of a product and firms have incentives to provide relevant information through advertising to maximize their profits. The other group of papers focuses on how advertising conveys "soft" (indirect) information, from which consumers can correctly infer unobserved quality of products. In these models, advertising which does not directly affect demand (neither persuasive nor containing explicit informative content) is called "dissipative" advertising. It is assumed that consumers can easily observe how much money has been spent; therefore the fact that a firm is "burning" money on advertising is enough to "signal" the firm's confidence in product quality to potential consumers.

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. Each firm makes an active choice of its advertising policy which consists of a choice of which price or prices to advertise and how many advertising to send out at each such price. Although the model presented in chapter 5 only focuses on firms' optimal advertising decisions and takes price as given, the framework used to study how advertising reaches consumers is grounded in Butters-type model. Therefore, I briefly review the most relevant parts of the Butters model below.

The random process by which the advertising is allocated can be pictured as the sellers dropping their advertising at random into buyer's mailboxes. Let r denote the total number of advertisements sent to buyers by a seller. Advertisements are assumed to be assigned independently with equal probability to each buyer. Suppose there are n buyers, and each buyer has a probability $\frac{1}{n}$ of receiving a given advertisement. A buyer thus can receive 0,1,2,... advertisements. For any given r and n, the probability that a given buyer receives x advertisements is given by the binomial distribution with r trials and probability $p = \frac{1}{n}$ of success for each trial. The binomial distribution approaches the Poisson distribution as n get large, holding $\frac{r}{n}$ fixed. Therefore, the probability that any given buyer receives x advertisements approaches $\exp\left(-\frac{r}{n}\right) \cdot \left(\frac{r}{n}\right)^x / x!$. A buyer's probability of not receiving any advertisement at all is $\exp\left(-\frac{r}{n}\right)$. Then the proportion of buyers who receive at least one advertisement and thus are reached by advertising is $\Phi = 1 - \exp\left(-\frac{r}{n}\right)$.

Grossman and Shapiro (1984) consider how advertising is used to inform consumers of products in the same way as Butters (1977) does, but allow product differentiation along two dimensions (both information and location). They extend Butters'

[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.

model to analyze oligopolistic interaction and allow firms to simultaneously choose price and advertising level. Although my work focus on analyzing firms' advertising decision in the context of monopolistic competition, however, similar to Grossman and Shapiro (1984), I consider both information differentiation and product differentiation (in terms of observed characteristics and unobserved quality).

Anderson and Renault (2006) extend the directly informative advertising literature by discussing firms' choices of the type of information transmitted in advertisements. Their model explains the lack of direct information in some advertisements by investigating the incentives for a firm to choose to advertise price only, match only, both price and match, or not to advertise at all. In order to consider the two dimensions of advertising: price and attributes, they assume consumers have to go to stores to check out the product prior to making purchases, therefore search cost is incurred. By using the models of consumer search which consider advertising as a means of transmitting information, they find that there is no incentive for firms to provide precise information on product characteristics only. The firm optimally uses minimum match advertising, which guarantees a threshold utility to the consumer and therefore reassures the consumer that visiting is desirable. Their results indicate that a forced disclosure policy for prices is not needed, and it is socially harmful to impose a full disclosure rule for product information if the firm can perfectly parse the information it conveys to the consumer.

Another explanation for the lack of information in advertisements stems from the "money-burning" theory of advertising. Nelson (1974) argues that the interaction of experience and repeat purchases can create an asymmetry in the returns to advertising and thus supports the signaling equilibrium. Since advertising costs are the same for all products with different quality level, therefore the returns to advertising must be greater for higher quality. This feature distinguishes an advertising signal from the type of signal discussed by Spence (1973) for which asymmetry in the cost of the signal, not the return, is fundamental. Kihlstrom and Riordan (1984) and, later, Milgrom and Roberts (1986) formalize Nelson's intuition with consumer rationality.

In the models of Kihlstrom and Riordan (1984), product quality is either High or Low. Firms are competitive price takers and decide whether to enter the high-quality market by spending some minimum amount on advertising as "entry fee", or, otherwise, enter the low quality market by default. Therefore, advertising is taken as simply a conspicuous expenditure of resources by firms and becomes part of firms' fixed cost. To discuss the conditions under which the advertising equilibrium exists, they set up two models with different assumptions about market information structure. In their model, however, firms do not choose prices. Instead, a firm's advertising alone determines whether customers believe it to be high or low quality, and once this assignment to one or the other submarket is made, prices are determined via a standard supply and demand model. In equilibrium, prices in fact end up being correlated with quality but are not used to infer quality.

In comparison with the perfect competition, Milgrom and Roberts (1986) assume that a firm with a new product is the sole producer and decide its introductory price and the level of dissipative advertising. In their model, both price and advertising may be used as signals for the initially unobserved quality of the newly introduced experience product. Repeat purchases play a crucial role in their model. Here I review the most relevant part of their model.

Let $\pi(p, q, \rho(p, a)) - a$ be the expected present value of the profits to a firm of true quality q ($q = H$ or L) who sets an introductory price of p, spends a on introductory advertising, and is believed with probability $\rho = \rho(p, a)$ to be producing quality H. They show that, at any separating equilibrium, the choice (p_H, a_H) of the high-quality firm must be a solution to

$$\max_{p,a} \pi(p, H, 1) - a$$

$$\begin{aligned} s.t. \pi(p, L, 1) - A &\leq \pi\left(p_L^L, L, 0\right), \\ p, a &\geq 0 \end{aligned}$$

In equilibrium, both advertising and price may be used as signals, with the chosen levels of prices and advertising differing between high and low quality firms. The extent to which each is used depends, in a rather complicated way, on the difference in costs across qualities. If the marginal cost of production is lower for high quality products, then price may be inversely correlated to quality and used as a signal for product quality. If the marginal cost is unrelated to product quality, then only advertising expenditures would be correlated to quality and used as a quality signal. Note that advertising here has no direct impact on demand or gross profits. Its only possible influence is through pre-purchase perceptions of quality. It is thus a purely dissipative signal. If actual quality were known by potential customers before purchase, then $\pi(p, q, q) - a$ would be the relevant profit function net of advertising expenditure for a firm known to be producing quality q. Clearly, the optimal advertising budget in these circumstances is $a = 0$.

Hertzendorf (1993) extends the analysis by relaxing the assumption that consumers can perfectly observe the firm's advertising expenditure. Instead, he takes into account the possibility that consumers observe the monopolist's advertising expenditure with error. The author shows that, under a reasonable condition, price and advertising expenditures would never be simultaneously employed to signal quality.

A strategy for the monopolist is a function, $M(q) : [H, L] \to (p, a)$, that translates the actual quality of the firm into a nonnegative price-advertising pair. In particular, the beliefs of all consumers are described by $Eq(p, \widetilde{a}) : (R_{++}xN) \rightarrow [0, 1]$, which translates an observed price-advertising pair into an expectation of quality. The level of advertising observed by consumers will not, in general, be the same as the level of advertising purchased by the monopolist. Suppose a_q is the advertisement purchased by type q monopolist (in equilibrium) and $g(\widetilde{a} : a_q)$ is the probability density function of observed advertisements by consumers. Then

$$E\pi(p, a_q, q, Eq(p, \widetilde{a})) = \sum_{\widetilde{a}=a_{\min}}^{a_{\max}} g(\widetilde{a} : a_q)\,\pi(p, a_q, q, Eq(p, \widetilde{a}))$$

demonstrates the importance of treating advertising as a stochastic process. In contrast to Milgrom and Roberts (1986), it shows that such a process will lead to pricing that is uncorrelated to quality. The basic idea is that whenever price is a sufficient statistic for quality, advertising expenditures are unnecessary and wasteful. The result shows that advertising will only be used to signal quality when there are no concurrent price signals from the monopolist. There is no mechanism by which the high quality monopolist can minimize the cost of signaling through simultaneous advertising and price signals.

My dissertation is inspired by those theories; however, it presents a model from empirical perspectives. The "money burning" signaling theory mainly focuses on high quality firms “burn” money to show their confidence to recoup their investments, therefore advertisements need no contents. However, 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 dissertation. 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.

2.2 Empirical Literature of Informative Advertising

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 two 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. They also find that models with persistent consumer uncertainty and learning fit data better.

Ackerberg (2001) empirically distinguishes two main different effects of advertising in nondurable, experience-goods markets: the informative effects and prestige effect of advertising. The basic idea to distinguish those two effects is very intuitive: advertising that informs consumers of a brand's inherent characteristics (include search characteristics and/or experience characteristics) should primarily affect inexperienced consumers (those who have not purchased the brand in the past); on the other hand, prestige or image effects of advertising should affect both inexperienced and experienced consumers relatively equally.

In this paper, Ackerberg uses both consumer-level data on purchases of a newly introduced brand of yogurt and consumer advertising exposures data over time. Empirical results indicate a significant effect of advertising on inexperienced consumers and either an insignificant or declining effect on experienced consumers. This study thus concludes that these advertisements for a newly introduced brand of yogurt are influencing consumer behavior primarily by informing them about search and experience characteristics, not by creating prestige or associating the product with favorable images. To check the robustness of the results, Ackerberg experiments with different types of reduced form models and finds that this differential effect of advertising on the behavior of experienced and inexperienced consumers is very robust over different models.

Those papers use reduced-form models to allow more flexibility and have less computational burden, however, they have some possible problems from which most reduced form models suffer. For example, if one believes optimal firm behavior implies a dynamic optimization problem, when a firm's current decision affects future states of knowledge, then these reduced-form models are approximations to the optimal dynamic decision rules. In that case, the quality of the results of those reduced-form models relies on the quality of those approximations. One disadvantage with reduced-form analysis is that it is hard to distinguish and separately quantify different effects of advertising. Another problem related to these reduced-form models is that they are unable to explicitly help answer important welfare questions about advertising. Therefore, there is a growing literature using structural models to empirically study different influences of advertising on consumer behavior and how firms make their optimal advertising choices.

Erdem and Keane (1996) develop a structural model of household information learning behavior in the laundry detergent market and analyze how advertising impacts consumer learning process. They use consumer level data of laundry detergent purchase decision and consumer level advertising exposure data from Nielson scan data. Since detergents are frequently and regularly purchased products, this paper use a dynamic model in which consumers learn about brand attributes in a Bayesian manner.

In their model, consumers who have imperfect information about a brand's attributes make purchase decisions each week, conditional on the information they get from the previous experience and advertising exposure. Therefore, consumer i's expected utility can be written as

$$E\left[U_{ij}\left(q_j\right)|I_i\left(t\right)\right] = E\left[U_{ij}\left(q_j\right)|A_{cijt}, Q_{eijt}\right]$$

q_j is the mean brand attribute level for brand j and is the same for all consumers. $I_i(t) = (A_{cijt}, Q_{eijt})$ contains all history of the information consumer i has obtained through her past consumption experience and/or advertising exposure experience. Each consumption experience provide a noisy signal $q_{eijt} = q_j + \delta_{ijt}$ about q_j and each advertising exposure also provides a noisy signal $a_{cijt} = q_j + \varsigma_{ijt}$ about q_j. Consumer i use these signals to update her expectations of brand attributes in a Bayesian manner. Over time, households have different experiences when they try brands, and they also may receive different advertising signals. Although consumers have the same priors about brands, their perception errors about mean brand attribute levels diverge over time as they receive different signals.

This paper provides a framework to analyze how consumers learn direct information about brand attributes from advertising. In fact, it discusses how the content of each advertising message provides noisy but direct information about brand attributes. Whenever consumer i watch at least one advertising in week t, she learns about q_j from the advertising content. Here, "direct" means consumers do not infer q_j from advertising spending or observed advertising intensity. Advertising intensity itself does not play any signaling role; instead, it reduces the variance of advertising messages. The more advertisement consumer i has received in the past weeks, the less uncertain she is about q_j.

Based on their 1996's paper, Erdem, Keane and Sun (2008) add price and advertising frequency/intensity as two more information sources for consumers. In this paper, consumer i's expected utility is written as

$$E\left[U_{ij}(q_j) \mid I_i(t)\right] = E\left[U_{ij}(q_j) \mid P_{ijt}, A_{Iijt}, A_{cijt}, Q_{eijt}\right]$$

$I_i(t) = (P_{ijt}, A_{Iijt}, A_{cijt}, Q_{eijt})$ contains all history of information consumer i has obtained through four information sources: price, observed advertising intensity, advertising exposure experience and consumption experience. Similar to Erdem and Keane (1996), both advertising exposure experience and consumption experience provide noisy but direct information about the mean brand attribute level for brand j. Both log value of price and transformed 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.

In this paper, advertising intensity not only reduces the variance of advertising messages, but also signals the product quality itself. The more advertisement consumer i has received in the past weeks, the less uncertain she is about q_j and the higher her perceived expected value of q_j is. By using Nielsen scanner data for ketchup category, they find that advertising frequency does influence consumer learning, although it is less quantitatively important than price. Their results also suggest that use experience is the most important signal of quality, followed by price, advertising frequency, and then advertising content. However, all four mechanisms appear to be important, because dropping any one of them led to a significant deterioration in model fit.

To fully explore the same question proposed in Ackerberg (2001), Ackerberg (2003) sets up a formal structural approach to formally model utility functions and information structures, taking the structural approach as a complement to the reduced-form approach. This paper uses a dynamic learning model that explicitly includes both informative and prestige effects of advertising. The identification argument used in Ackerberg(2003) is similar to the one used in Ackerberg (2001), from a more structural perspective. In this paper, consumer i's expected utility can be written as

$$E\left[U_{ij}\left(q_j\right)|I_i\left(t\right)\right] = E\left[U_{ij}\left(q_j, A_{Iijt}\right)|A_{Iijt}, Q_{eijt}\right]$$

where q_j is the mean experience utility consumer i obtain form the brand. Consumer i wants to learn about q_j either through her past experience Q_{eijt} which provides direct information on experience characteristics or observed advertising intensity A_{Iijt} which provide indirect, signaling information on experience characteristics. Therefore, observed advertising intensity by consumer i at week t is also assumed to be a linear function of q_j. In addition, observed advertising intensity A_{Iijt} also enters utility function directly to capture the possibility that advertising itself may provide additional utility to consumers through the prestige effect.

All three papers discussed here 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 third paper 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 q_j and analyze how consumers infer q_j from observed advertisement intensity. However, in this dissertation, advertisement intensity as a function of observable attributes and unobservable quality is solved as an equilibrium results.

In all three discussed papers, consumers are assumed to be aware of all products for sale. Sovinsky (2008) relaxes this assumption by estimating a model of limited consumer information, in which advertising influences the set of products from which consumers choose to purchase. In this paper, she investigates the U.S. personal computer market where top firms spend over $2 billion annually on advertising. In her model, advertising determines information set of consumers and provides hard information about product availability. The probability that consumer i purchases computer j can be written as

$$s_{ijt} = IS\left(a_{ijt}\right) \cdot pro\left(U\left(X_{ijt}\right) > k\right)$$

which is composed of the probability consumer i is informed by advertising($IS\left(a_{ijt}\right)$) and the probability the consumer's utility level is above certain threshold level.

Instead of considering only the impact of advertising on consumer behavior, Sovinsky (2008) explicitly models firms' behavior, which allows her to consider the optimal choices of advertising in light of the effects on demand. The supply side analysis also allows her to estimate the marginal cost and then calculate markups, so that she can compare her results with the traditional models with full information assumptions. She finds estimated markup is much higher than that predicted by full information models. The estimates indicate median markups would be 5% under full information, one-fourth the magnitude of those under limited information. 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, Sovinsky (2008) does not consider the possible uncertainties faced by consumers about product attributes, therefore s_{ijt} is a function of indirect utility $U\left(X\right)$ instead of expected utility. In my model, the probability that consumer i purchases computer j can be simplified as

$$s_{ijt} = IS\left(a_{jt}\right) \cdot pro\left(E\left[U\left(X_{ijt}, q_j\right) | a_{jt}\right] > k\right)$$

in which 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).

2.3 Related Literature about the Motion Picture Industry

The theatrical motion picture industry has been very appealing for academic researchers both in economics and marketing disciplines, because it provides rich data and interesting phenomenon for scholars to explore. There are some studies that investigate the impacts of advertising or/and social learning in this industry.

A number of empirical studies about advertising in the movie industry are mainly from the marketing literature. Most of them focus on studying the effect of advertising on box office revenue which is confounded by the classic endogeneity problem caused by movie's quality. Elberse and Anand (2007) pursue a different empirical strategy by using data from the Hollywood Stock Exchange to study the impact of movie advertising on a measure of sales expectations in the pre-release period. Their measure is the movie's "stock price" as it trades on the Hollywood Stock Exchange, a popular online stock market simulation. Besides investigating whether or not pre-release advertising affect the updating of market-wide expectations, they are also interested in studying whether this effect varies according to product quality. They use critical reviews as the measure of movie quality, which has the disadvantage that critics' views do not necessarily reflect the quality perceptions of the general public. Therefore, the measure represents a relevant dimension of quality. 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. In addition to their work, my dissertation further provides the mechanism which supports the positive impact of advertising on market-wide expectation and explains the interaction of advertising and movie quality.

Basuroy, Desai, and Talukdar (2006) empirically test the role of potential signals of movie quality in the motion picture industry using a reduced-form analysis method. This paper tests for the attenuating role of third-party information sources, such as critics' review consensus and cumulative word of mouth, on the strength of the two plausible signals: sequels and advertising expenditures. The basic intuition is that the presence of information from independent or third-party sources about a movie's quality should moderate the strength of the possible signals. Attenuation (lack of attenuation) of the positive effects of sequels and advertising expenditure on box office performance in the presence of independent information about movie quality (e.g., consensus among critics' reviews) would be consistent (inconsistent) with their potential signaling roles. This paper analyzes the data with a dynamic simultaneous-equations model. Specifically, the authors construct a system of three interdependent equations with revenues, advertising expenditure and screens as the dependent variable separately. They find evidence that is consistent with their hypotheses about the potential signaling roles of sequel and ad expenditure both at the release phase across movies and over the post-release phases for any movie, and about the positive interaction between sequels and ad expenditure.

This paper focuses on providing a diagnostic empirical test for potential signals of quality in the movie industry. Their results directly support my model which emphasizes on the signaling role of advertising. Although this paper tries to test whether there is an attenuation effect on box office revenue with cumulative word-of-mouth, the authors use the cumulative number of screens (since a movie's release) to approximate the cumulative level of word-of-mouth communication for a movie, which is problematic itself. Due to the limitations of reduced form analysis, it is impossible for them to analyze how advertising affects demand and how firms make their decisions. In contrast, the structural approach I am using in this dissertation allows me to differentiate the effects of advertising and analyze consumers' and firms' choices in detail.

Another set of papers mainly focus on identifying the impact of social learning on a movie's box office performance. There is a large and influential theoretical literature on the topic of social learning, while the empirical evidence is limited. Most of the existing empirical evidence is from a growing number of studies based on laboratory experiments. Although laboratory experiments are useful, real-world data are necessary to establish how important social learning is in practice. Moul (2007), Moretti (2009) and Santugini (2007) are among the first studies to credibly test for social learning using real-world, industry-wide data.

Moul (2007) tries to identify and measure the impact of word of mouth on U.S. theatrical movie admissions. This paper incorporates a static discrete choice model for consumer behavior, and derived movie market shares are expressed as standard logit probabilities. The word-of-mouth effect on demand is modeled as a part of the error term, since it affects consumer behavior but is unobservable by the econometrician. The author uses the correlation of the error terms for a particular movie across time to identify and measure the WOM. He finds that information about movie quality travels quickly among consumers. This paper presents evidence of informational externalities in the motion picture market that have a significant impact on consumer behavior.

Although this paper tries to address the issue of information transmission through WOM in the movie theatrical market, the author doesn't explicitly model how information transmission through WOM impact consumers' choices. In fact, his model is inconsistent with information learning story, since consumers' utility is a function of observable characteristics of movies without uncertainty (observable to consumers but unobservable to econometricians) and consumers know all components of the error term. The author uses a reduced form model to analyze the impact of WOM on demand, by writing the error term as a function of past cumulative admissions and a movie's age in theaters.

Moretti (2011) also tries to identify and quantify the impact of social learning on consumer demand, but use different empirical strategies. Moretti defines social learning as the process through which individuals use feedback from their peers to update their own expectations of movie quality. The author first sets up a simple theoretical model to describe the consumer information learning in the Bayesian learning framework, then generates some transparent and testable predictions to bring to the data, and use reduced-form models to empirically test the predictions which are consistent with social learning story.

The key empirical identification strategy used in this paper is to compare the change in sales for movies over time with a positive and a negative surprise. The surprise is defined as the difference between realized box-office sales and predicted box-office sales in the opening weekend. Specifically, the author uses the residual from a regression of first-week log sales on log number of screens as his measure of movie-specific surprise. The estimates suggest that social learning appears to be an important determinant of sales in the movie industry, accounting for 32% of sales for the typical movie with positive surprise.

Santugini (2007) departs from Moul (2006) and Moretti (2011) by incorporating forward-looking behavior and observation learning, as well as proposing a new approach to account for demand saturation. He derives the demand for movies using a dynamic discrete choice model in which consumers are endowed with private information about a movie and engage in as well as anticipate learning. He measures the extent to which consumers learn about the quality of a movie from observing its market share in the release week. He also assumes that consumers watch a movie at most once to account for demand saturation; therefore he endogenizes the consumer base for a particular movie.

After controlling for demand saturation and movie competition in the theater, Santugini finds evidence of observational learning by using the variation in the market share growth rate across weeks. Given the estimated parameters, he also measures the effect of observational learning on movie demand in the week after the release week and the effect of anticipation of observational learning on demand for a movie in its release week by running different counterfactual experiments.

Although social learning and advertising are believed to be two key factors faced by both the demand and the supply side in the motion picture industry, very few empirical studies has investigated the interaction of those two factors and how both demand and supply side take those two factors into account when making their decisions. Joo (2009) is particularly related to my work in terms of investigating how social learning can impact the effectiveness of movie advertising. She develops an equilibrium model describing consumer decisions about watching a film as well as movie distributor decisions about how much to spend on advertising at the time of the movie's release.

Consumers are assumed to have uncertainty about movie quality. The indicator of movie quality, which is also called the common belief of movie quality, enters consumers' utility function as a part of the observed characteristics of a movie. Information about movie quality is initially based on critics' ratings, and the prior distribution for a movie's quality is the same as the distribution of its initial critics' ratings. For critics' ratings, the author uses the Metascores available through Metacritic (www.metacritic.com). The posterior distribution is obtained by Bayesian updating, using the number of previous moviegoers and their ratings reported over the Internet. The mean and variance of reported rating on Internet Movie Database

(IMDB) is used to measure the true distribution of quality perceived by previous moviegoers. The speed of information update is fixed at a reasonable value instead of being estimated from data in this paper. Although I also use Bayesian updating framework to study consumer learning, I depart from Joo (2009) by estimating the prior and posterior distribution of the movie quality and recovering each movie's true quality from both the market performance and advertising data. More importantly, the speed of information update is not fixed but one of important parameters to be estimated in my study.

In Joo (2009), the total advertising spending is determined by the studio's profit maximization based on the expected movie demand. The total advertising spending on the theatrical market is taken as an indicator for the effect of advertising in the opening week and enters the utility function as one of the observed movie characteristics. Therefore, advertising affects consumers' utility directly and is taken as the predominant way to boost a movie's demand. Movie producers make a single time decision for advertising expenditure, but they fully consider the expected future revenue stream of the movie. Therefore the advertising campaigns are given before a film's release and the duration of a movie is both given and known. (Since the optimal time to discontinue a movie's run is affected by the arrival of new information, it is not good ex-ante to commit to a fixed termination date.) I depart from Joo (2009) by focusing on the informative roles of advertising and analyzing how the interaction of advertising and social learning helps solving the information asymmetry and uncertainty problems in the movie industry. In my model, I assume advertising affects consumers' information set instead of affecting utility directly and firms decide the allocation of advertising spending over time instead of assuming firms make a single time decision for advertising spending. Therefore the observed large proportion of a movie's advertising budget used in the weeks leading up to its theater debut is an endogenous choice in my model instead of an assumption, and it is consistent with the idea that advertising signaling has occurred prior to product entry.

Joo (2009) finds that the social learning channel can amplify or impede the effectiveness of advertising, depending on the quality of a movie and the degree of uncertainty about quality. The simulation results show that for good movies, producers spend 4.2% more on advertising with learning than they would without learning. For bad movies, social learning makes a 1.4% increase in the level of advertising expenditure. Uncertainty about movie quality causes studios to spend 3.6% more on advertising for good movies, but 1.4% more for bad films. Although this paper does not take into account that studios' decisions are informative to consumers, its results presents evidence that studios' advertising strategies are highly related to their movies' quality and therefore should provide information about movie quality to consumers. In my model, studios' decisions are informative to consumers, and studios anticipate consumers to learn from their decisions, which, in turn, affects studio behavior.

3 Description of Data and Industry

3.1 Industry Background

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 million, and studios spent another $35.9 million on marketing that movie to the public (www.mpaa.org) [4].

New movies enter the theater market every week and exit the market after a few

[4] MPAA figures include not only buying ads, such as commercials on TV and pages in newspaper, but also costs of publicity, movie trailers, and creating marketing materials.

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 the 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 reflects studios' eagerness to emphasis 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[5] 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. Market research is used to get more information, allocate advertising resources and manage risks. To evaluate the movie's playability and marketability, studios usually hire outside research vendors to run the test screening. Several groups of audiences are chosen from the targeted moviegoing consumer segments and exposed to the nearly finished or already finished movies. After watching a movie, audiences are asked about what they like and don't like,

[5] In this dissertation, I use "studios" and "distributors" interchangeably. In practice, major studios do play the role of distributor, while independent distributors tend to fill marget segments which are not covered by majors.

including all aspects of the movie. Data drawn from this small slice of the potential consumer are then projected to a larger population. (Marich, 2009) Tracking surveys which start six weeks before theatrical release are used to quantify consumer awareness of movies. Movie studios may change their advertising spending by using the information from the tracking surveys. At the end of the six-week tracking arc, the tracking results are also used to predict the opening-weekend box-office revenue.

After the release of the movie, the movie studio can elicit more information about its movie's unobserved quality from the market responses: the box office performance in the opening weekend, requests from exhibitors for increasing or decreasing the number of screens, and, the most important, comments from consumers who actually watched the movie. One way to collect information is to do exit surveys at theater locations to interview moviegoers right after they have watched a movie. Respondents of the exit survey are better representatives of the potential target audiences and provide more accurate demographic information which is a key input when studios make their marketing decisions. For example, if the exit survey shows that more ticket-buying adults for an animation movie actually did not come with children, then the marketing effort may need to targeting adults without children more. Another key question to ask is whether the respondent will recommend the movie to peers. Answers to that question can help studios to decide how heavy or light future waves of advertising should be. Marich (2009) provides more details about pre-release and post-release market research.

Those 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[6] with their advertising budget. In my data sample, total advertising spending aver-

[6] "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.

ages around 50% of the production budget, while the ratio between those two varies widely across movies. It ranges from 1.5% to 875%, with a standard deviation of 65.5%. Figure 1 further shows that although the pre-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. 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. Big movies with big marketing campaigns bomb all the time for several reasons. First, it's increasingly difficult to recruit test audiences that are representative of the moviegoing population. Second, market research for movies focuses on the regular moviegoer population segment, and, therefore, may miss some unexpected audience segments. For example, animation movies such as "Wall-E" (2008) and "UP" (2009) may unexpectedly attract the adult audience. Third, mistakes in processing raw audience data may happen when rapid evaluation is needed to meet deadlines of movie companies.

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

3.2 Data

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"[7] 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.).[8] Advertising spending over movie's 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

[7] 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.

[8] I developed an excel application using VBA (Visual Basic for Application) that can connect the website and download the data automatically. All the downloaded data were stored to a local Access database.

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[9]. 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 ac-

[9] My 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.

count. Still, I include the number of other movies released widely in the same week to control for competitive 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. households 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.

3.3 Nonstructural Analysis

Before I conduct the structural analysis, I perform a non-structural analysis to understand intuitively what factors determine a movie's advertising spending and its performance over time. Non-structural analysis also provides insights about advertising's signaling effect in the per-release stage and social learning's influence in post-release stage.

I choose two sub-samples from my dataset: one sample includes 632 movies which have no prequels, and the other includes 111 sequel movies. For sequel movies, the information asymmetry between distributors and moviegoers ought to be lower before the release, compared to non-sequel movies; the signaling effect of advertising and social learning should be less important in determining a sequel movie's market performance if those factors do play a role.

For simplicity, I assume the distributor of movie j decides its optimal advertising spending for the per-release period (period 1) and post-release period (period 2) in the following way:

$$\begin{aligned} a_{jt} &= A\left(x_{jt}, q_{jt}\right) \\ &= \gamma_t x_{jt} + f\left(q_{jt}\right) \end{aligned}$$

Here, x_{jt} contains observed characteristics in period t and q_{jt} presents the distributor's information about the movie's unobserved quality. $q_{j1} = q_{js}$ means that the distributor decides its advertising spending for period 1 according its private information about its movie's quality. $q_{j2} = q_j$ means the distributor gets accurate information about its movie's quality and then decides its advertising spending for period 2. $f\left(q_{jt}\right)$ is an increasing function of q_{jt}, which means the distributor spends more on advertising if it is more confident about the movie's quality. Since q_{jt} is unobservable for me as an econometrician, $f\left(q_{it}\right)$ will be in the error term if I run linear regression of a_{jt} on x_{jt}. The number of box office tickets sold in period t is modeled as:

$$\begin{aligned} bot_{jt} &= \varphi\left(\sum_{s=1}^{t} a_{js}\right) \tau_t\left(x_{jt}, E_t\left(q_j\right)\right) \\ &= \exp\left(\lambda_{0t}\right)\left(\sum_{s=1}^{t} a_{js}\right)^{\lambda_{1t}} \exp\left(\beta_t x_{jt} + \mu_t E_t\left(q_j\right)\right) \end{aligned}$$

By taking the logarithm of both sides, I can estimate a log-linear version of above equation:

$$\begin{aligned} \ln\left(bot_{j1}\right) &= \lambda_{01} + \lambda_{11} \ln\left(a_{j1}\right) + \beta_1 x_{j1} + \mu_1 E_1\left(q_j\right) + \epsilon_{j1} \\ \ln\left(bot_{j2}\right) &= \lambda_{02} + \lambda_{12} \ln\left(a_{j1} + a_{j2}\right) + \beta_2 x_{j2} + \mu_2 E_2\left(q_j\right) + \epsilon_{j2} \end{aligned}$$

Since $E_1\left(q_j\right)$ and $E_2\left(q_j\right)$ are unobserved, I use the residuals from linear regression of

a_{it} on x_{jt} to approximate them.

Table 2 and Table 3 show the estimation results for both advertising spending and box office performance in each period. When comparing the coefficients for advertising spending in Table 3 of those two types of movies in both periods, we can see that sequel movies have higher significantly positive coefficients for advertising spending than non-sequel movies. This may imply that advertising spending improves box office performance by reaching more potential consumers, and it reaches consumers more effectively when consumers are already familiar with the movie's concepts and characters. The more interesting results are the coefficients of $f(q_{js})$ and $f(q_j)$ which are residual terms from advertising spending regressions and used as proxies for $E_1(q_j)$ and $E_2(q_j)$. In period 1, the coefficient of $f(q_{js})$ is significantly positive for non-sequel movies, but negative (only significant on 10% level) for sequel movies. This result shows us that, when there is greater information asymmetry between distributors and consumers, quality information conveyed (or signaled) by advertising spending is more important for consumers' movie-going decisions. In period 2, the coefficient of $f(q_j)$ is significantly positive for non-sequel movies and not significant for sequel movies, which also shows that information learning is more important for movies with greater information uncertainty and asymmetry. Meanwhile, conditional on box office performance in period 1 and $f(q_j)$, the coefficients of $f(q_{js})$ for both types of movies in period 2 are significantly negative, which may represent consumers' revenge as a response to false signals.

Although this is a simplified data analysis, it still shows that information uncertainty and learning are very important factors in determining a movie's box office success. When levels of information asymmetry differ between distributors and consumers, information learning affects movie demand to varied extents. Therefore, a structural model is necessary, to fully understand and measure how advertising and

WOM through social learning can help reduce the information uncertainty and asymmetry in this industry.

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 theatrical market. 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. In what follows, I will suppress the movie index j for notational simplicity.

Period 0 (after a movie is produced): a single studio produces a new movie with observable attributes x and unobservable quality q. Instead of knowing q perfectly, the studio receives a noisy signal of its movie's quality, q_s. 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 getting to know the availability of the movie, 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 who may enter the market at 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 will 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, is a random draw from its population distribution $\widetilde{q} \sim iidN\left(\overline{q}, \sigma_q^2\right)$.

4.2.1 Information Learning: Studio

As discussed in Chapter 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_s, of the movie's true quality q in period 0. $q_s = q + \varepsilon_s$, with $\varepsilon_s \sim iidN\left(0, \sigma_s^2\right)$, is known only by the studio. Here, σ_s^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_s to update its prior expectation of q according to the Bayesian updating rule:

$$E_1^s\left(q\right) = E^s\left[q|q_s\right] = \overline{q} + \beta_q\left(0\right)\left(q_s - \overline{q}\right)$$

where $\beta_q^s\left(0\right) = \frac{\sigma_q^2(0)}{\sigma_q^2(0)+\sigma_s^2} = \frac{\sigma_q^2}{\sigma_q^2+\sigma_s^2}$ is the weight the studio puts on the noisy signal q_s. $E_1^s\left(q\right)$ is the weighted average of prior expected value of q and the noisy signal q_s. When the signal is more accurate (with smaller value of σ_s^2), more weight should be put on the signal received by the studio. The perception variance by the studio for period 1 is given by

$$\sigma_q^{s2}\left(1\right) = \frac{1}{\frac{1}{\sigma_q^2(0)} + \frac{1}{\sigma_s^2}} = \frac{\sigma_q^2\sigma_s^2}{\sigma_q^2 + \sigma_s^2} = \sigma_q^2\left(\frac{\sigma_s^2}{\sigma_q^2 + \sigma_s^2}\right)$$

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 $f\left(q_s \mid q\right)$ satisfies the Monotone Likelihood Ratio Property (MLRP) in q, for any value q^*, $\Pr\left(q \geq q^* \mid q_s^{'}\right) \geq \left(q \geq q^* \mid q_s\right)$ if $q_s^{'} > q_s$. Intuitively, the better the received signal q_s is, the probability that the movie has quality above certain level is higher and the value of $E_1^s\left(q\right)$ is higher. The studio then decides its optimal advertising spending a_1 according to its perceived movie quality $E_1^s(q)$.

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 is revealed to the studio at the end of Period 1 and the studio adjusts its optimal advertising spending a_2 for period 2 according to q.

4.2.2 Information Learning: Consumers

In this section, I will 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".

In practice, movie studios tend to hire outside media-buying advertising agencies to handle purchases of advertising. When they map out a plan for media buying, they emphasize two important performance metrics: reach and frequency. Reach refers to the percentage of households or population in a target that see an advertisement at least once in a measurement period. It is a measure of the breadth of an advertising campaign. Frequency refers to a percentage that expresses the number of times households or persons in a target audience are exposed on average to advertisements in a measurement period. It is a measure of depth of an advertising campaign. (Marich 2009) With the help of independent measurement companies, studios have an estimate of those two performance measurements in advance when making their advertising decisions. Therefore, it is reasonable to consider how many consumers can be reached and how many advertisements a consumer can receive when modeling the impact of advertising on demand.

Suppose that, at the beginning of period 1, the optimal advertising spending the studio decides is a_1, and then advertisements are sent out to consumers as a series of messages. A consumer can receive 0, 1, 2, $\cdots$ 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_{i1} advertisements in period 1 is given by the binomial distribution which approaches the Poisson distribution. k_{i1} is the realized advertisement intensity observed by consumer i from $\widetilde{k}_{i1} \sim pois(\lambda a_1)$. Here, λ 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

$$\begin{aligned} \varphi_1(a_1) &= prob(k_{i1} > 0) = 1 - prob(k_{i1} = 0) \\ &= 1 - \exp(-\lambda a_1) \end{aligned}$$

which is also the market coverage rate.

In the post-release weeks, the studio adjusts its optimal advertising spending to a_2, 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:

$$\begin{aligned}\varphi_2(a_1, a_2) &= \varphi_1(\mu a_1) + (1 - \varphi_1(\mu a_1))\, prob\,(k_{i2} > 0) \\ &= [1 - \exp(-\lambda(\mu a_1 + a_2))]\end{aligned}$$

Here, μ 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.

The assumption that consumers learn about the movie's arrival and its observable attributes only through advertising in both period seems a bit strong, considering that consumers may also learn about the existence of a movie through WOM. However, I believe it is still a reasonable assumption and does not affect the main conclusions of this paper for the following reasons. First, WOM in both periods can be taken as being induced by advertising. Therefore, whenever consumers are reached by WOM, they are just indirectly reached by advertising. Consumers can learn about the existence of a movie by WOM in both periods. While, the main conclusions of this paper are based on the key feature that consumers can only learn about the unobservable quality through firm-generated information (advertising) in period 1 and then through consumer-generated information (WOM) in period 2. Second, advertising (movie trailers) can convey much more information about observable attributes of a movie than WOM, not just the existence of a movie. Therefore, I assume that consumers are only aware of the movie when receiving advertisements, enter the market and make their watching decisions.

Here I assume that consumers are fully rational, which means consumers can 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 $a_1 = A(q_s)$ is an increasing function of received noisy signal, q_s. Intuitively, when the studio receives a better signal q_s, it spends more on advertising, and consumers tend to observe higher advertising intensity k_{i1}. Therefore, I assume that a rational consumer should take k_{i1} as a noisy signal of a_1 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 for consumer i in period 1 after observing advertisement density k_{i1} are

$$g_{i1}(q|k_{i1}, A_1(q_s)) = \frac{\int f(k_{i1}|A_1(q_s))\, f(q_s \mid q)\, d_{q_s|q} g_0(q)}{\int \int f(k_{i1}|A_1(q_s))\, f(q_s \mid q)\, d_{q_s|q} g_0(q)\, d_q}$$

$$E^c_{i1}[q \mid k_{i1}, A_1(q_s)] = \int q \cdot g_{i1}(q|k_{i1}, A_1(q_s))\, d_q \tag{1}$$

The Role of WOM When consumers enter the market at different time, they should have different information sources regarding to a movie's quality. For those who enter the market in the opening week, advertising is the main information source they can use to infer a movie's quality. While for those who enter the market in post-release weeks, they enjoy the extra benefit of information from 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 j talks to consumer i about the movie's quality, consumer i receives a WOM signal q_{iw} which is a random drawn from distribution $iidN\left(q, \sigma_w^2\right)$, where q is the true quality of the movie and σ_w^2 is the variance of the WOM signal. Consumer i may get several WOM signals and aggregate all information. Let δ_1 be the number of tickets sold in period 1, which is also the movie's box office performance in the opening week. ρ is the average proportion of consumers who like to share their movie-watching experiences with the representative consumer i. Here, ρ measures the information transmission speed, and the higher value of ρ implies that more consumers like to spread information through WOM. Therefore, consumer i gets a sample mean of experience signals, $\overline{q}_{iw}$, which is a random drawn from distribution $iidN\left(q, \frac{\sigma_w^2}{\rho\delta_1}\right)$. The variance of $\overline{q}_{iw}$ is a decreasing function of δ_1, which means that the more consumers watch the movie in period 1, the more accurate the average WOM signal, $\overline{q}_{iw}$, is about the movie's true quality, q.

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, δ_1, in period 1. With the existence of a signaling equilibrium, $\delta_1 = \Delta_1\left(A_1\left(q_s\right)\right)$ should be an increasing function of q_s and a rational consumer should infer q_s from δ_1. 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 for consumer i in period 2 is

$$g_{i2}\left(q|\overline{q}_{iw}, \delta_1, A_1\left(q_s\right)\right) = \frac{f\left(\overline{q}_{iw} \mid q, \frac{\sigma_w^2}{\rho\delta_1}\right) f\left(\Delta_1\left(q_s\right) \mid q, A_1\left(q_s\right)\right) g_0\left(q\right)}{\int f\left(\overline{q}_{iw} \mid q, \frac{\sigma_w^2}{\rho\delta_1}\right) f\left(\Delta_1\left(q_s\right) \mid q, A_1\left(q_s\right)\right) g_0\left(q\right) dq}$$

$$E^c_{i2}\left[q \mid \overline{q}_{iw}, \delta_1, A_1\left(q_s\right)\right] = \int q \cdot g_{i2}\left(q|\overline{q}_{iw}, \delta_1, A_1\left(q_s\right)\right) dq \tag{2}$$

If consumer i infers q_s from δ_1, then the prior expected q for consumer i is $E^c_{i1}\left[q \mid \delta_1\right] = E^s_1\left(q\right) = \overline{q} + \beta_q\left(0\right)\left(\Delta_1^{-1}\left(\delta_1\right) - \overline{q}\right)$, and equation (2) can be written as simple as

$$E^c_{i2}\left(q\right) = E^c_{i1}\left[q \mid \delta_1\right] + \beta^c_q\left(1\right)\left(\overline{q}_{iw} - E^c_{i1}\left[q \mid \delta_1\right]\right)$$

where $\beta^c_q\left(1\right) = \frac{\sigma^{c2}_q(1)}{\sigma^{c2}_q(1)+\frac{\sigma^2_w}{\rho\delta_1}} = \frac{\sigma^{s2}_q(1)}{\sigma^{s2}_q(1)+\frac{\sigma^2_w}{\rho\delta_1}}$ is the weight that consumer i put on the WOM signal $\overline{q}_{iw}$. The perception variance by consumer i for period 2 is

$$\sigma^{c2}_q\left(2\right) = \frac{1}{\frac{1}{\sigma^{c2}_q(1)} + \frac{\rho\delta_1}{\sigma^2_w}} = \frac{1}{\frac{1}{\sigma^{s2}_q(1)} + \frac{\rho\delta_1}{\sigma^2_w}} = \frac{\sigma^2_w}{\rho\delta_1}\beta^c_q\left(1\right)$$

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 this movie conditional on they are reached by advertising. Consumer i's expected utility from watching movie j in period t ($t = 1,\ 2$) is (the subscript j will be dropped for notational ease):

$$Eu_{it} = \gamma x + \alpha TD + E_{it}^{c}\left[q|I_i\left(t\right)\right] - p + \eta_t + \varepsilon_{it}$$

Here, x is composed of observed characteristics of the movie, such as genre, production budget, studio, the MPAA rating, the holiday indicator, etc., and γ is composed of consumer taste parameters. TD is the time dummy which indicates whether it is period 2 or not and α is the utility weight that consumer i attaches to TD. $E_{it}^{c}\left[q|I_i\left(t\right)\right]$ is the expected quality of the movie perceived by consumer i conditional on her information set, $I_i\left(t\right)$ in period t, and p is the price of watching a movie in the theater which is assumed the same for movies of different quality levels. η_t is the realized aggregate demand shock in period t from $\widetilde{\eta}_t \sim iidN\left(0, \sigma_{\eta_t}^2\right)$, and ε_{it} is consumer i's realized idiosyncratic preference shock in period t from $\widetilde{\varepsilon}_{it} \sim iidEV$. Consumer i's utility from the outside option in week t is $u_{it} = \varepsilon_{i0t}$ with mean utility normalized to zero, and $\widetilde{\varepsilon}_{i0t} \sim iidEV$.

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:

$$w_{i1} = w_{i1}\left(k_{i1}; A_1\left(q_s\right)\right) = \begin{cases} 1 & if \quad Eu_{i1} \geq 0 \\ 0 & if \quad Eu_{i1} < 0 \end{cases}, \tag{3}$$

and her watching decision in period 2 is described as:

$$w_{i2} = w_{i2}\left(\overline{q}_{iw}, \delta_1; A_1\left(q_s\right)\right) = \begin{cases} 1 & if \quad Eu_{i2} \geq 0 \\ 0 & if \quad Eu_{i2} < 0 \end{cases}. \tag{4}$$

Both equations (3) and (4) tell us that an informed consumer i will watch the movie in period t only when $Eu_{it} = \gamma x + \alpha TD + E_{it}^c\left[q|I_i\left(t\right)\right] - p + \eta_t + \varepsilon_{it} \geq 0$. Then the probability that the informed consumer i chooses to watch the movie in period t is $\tau_{it}\left(x, E\left(q|I_i\left(t\right)\right), \eta_t\right) = \Pr\left(Eu_{it} > Eu_{i0t}\right) = \frac{\exp\left(\gamma x+\alpha TD+E_{it}^c[q|I_i(t)]-p+\eta_t\right)}{1+\exp\left(\gamma x+\alpha TD+E_{it}^c[q|I_i(t)]-p+\eta_t\right)}$. In period 1, the probability that consumer i chooses to watch the movie, conditional on observing $k_{i1} > 0$ is given by

$$\begin{aligned} \tau_{i1} &= \frac{\exp\left(\gamma x + \alpha TD + E_{i1}^c\left[q \mid k_{i1}, A_1\left(q_s\right)\right] - p + \eta_1\right)}{1 + \exp\left(\gamma x + \alpha TD + E_{i1}^c\left[q \mid k_{i1}, A_1\left(q_s\right)\right] - p + \eta_1\right)} \\ &= \tau_{i1}\left(k_{i1}, \eta_1; x, A_1\left(q_s\right)\right) \end{aligned}$$

With advertising spending a_1 and market size M, the number of tickets sold in period 1 is:

$$\begin{aligned} \delta_1 &= \varphi_1\left(a_1\right) M \sum_{k_{i1}=1}^{\infty} \tau_{i1}\left(k_{i1}, \eta_1; x, A_1\left(q_s\right)\right) f\left(k_{i1}|a_1\right) \\ &= \varphi_1\left(a_1\right) M \tau_1\left(a_1, \eta_1; x, A_1\left(q_s\right)\right) \\ &= \Delta_1\left(a_1, \eta_1; x, A_1\left(q_s\right)\right) \end{aligned} \tag{5}$$

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

$$\begin{aligned}\tau_{i2} &= \frac{\exp\left(\gamma x + \alpha TD + E^c_{i2}\left[q \mid \overline{q}_{iw}, \delta_1, A_1\left(q_s\right)\right] - p + \eta_2\right)}{1 + \exp\left(\gamma x + \alpha TD + E^c_{i2}\left[q \mid \overline{q}_{iw}, \delta_1, A_1\left(q_s\right)\right] - p + \eta_2\right)} \\ &= \tau_{i2}\left(\overline{q}_{iw}, \delta_1, \eta_2; x, A_1\left(q_s\right)\right)\end{aligned}$$

With advertising spending a_2, the number of tickets sold in period 2 is:

$$\begin{aligned}\delta_2 &= \left[\varphi_2\left(a_1, a_2\right) M - \delta_1\right] \int \tau_{i2}\left(\overline{q}_{iw}, \delta_1, \eta_2; x, A_1\left(q_s\right)\right) \phi\left(\overline{q}_{iw} \mid q, \frac{\sigma^2_w}{\rho \delta_1}\right) d_{\overline{q}_{iw}} \qquad (6) \\ &= \left[\varphi_2\left(a_1, a_2\right) M - \delta_1\right] \tau_2\left(q, \eta_2; \delta_1, x, A_1\left(q_s\right)\right) \\ &= \Delta_2\left(a_2, q, \eta_2; a_1, \delta_1, x, A_1\left(q_s\right)\right)\end{aligned}$$

where $\left[\varphi_2\left(a_1, a_2\right) M - \delta_1\right]$ is the set 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 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_t in period $t-1$ for advertising campaigns in period t for $t = 1$, 2, based on its information about the movie's quality. I denote $S_t = \left\{E^s\left[q|I\left(t\right)\right], \sigma^{s2}_q\left(t\right), x, \delta_{t-1}, a_{t-1}\right\}$, as the set of state variables that are relevant to the decision of the studio. The per period expected profit for the studio in period

1 and 2 are

$$\pi_1(a_1; S_1) = \int [\Delta_1(\eta_1; S_1, a_1) p - a_1] d_{\Phi(\eta_1)}$$

$$\pi_2(a_2; S_2) = \int [\Delta_2(\eta_2; S_2, a_2) p - a_2] d_{\Phi(\eta_2)}$$

Then the value function for the studio is

$$\begin{aligned} V(S_1) &= \max_{a_1 \geq 0} [\pi_1(a_1; S_1) + E_1 V(S_2|a_1, S_1)] \\ V(S_2) &= \max_{a_{21} \geq 0} [\pi_2(a_2; S_2)] \end{aligned}$$

where $E_1 V(S_2|a_1, S_1) = \int V(q|a_1, \delta_1, S_1) d_{\Phi^s(q|I(1))}$ and $q|I(1)$ follows normal distribution with mean $E_1^s(q)$ and variance $\sigma_q^{s2}(1)$. It should be noted that the studio explicitly takes into account the effect of its advertising decision a_1 on the next period's expected mean quality $E_{i2}^c(q)$ and variance $\sigma_q^{c2}(2)$ perceived by consumers through opening week market performance δ_1.

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

$$\begin{aligned} a_2^* &= \max\left(\frac{\ln(\lambda M p E_2(\tau_2(q, \eta_2; \delta_1, x, A_1(q_s))))}{\lambda} - \mu a_1, 0\right) \\ &= A_2(q; a_1, \delta_1, x, A_1(q_s)) \end{aligned} \tag{7}$$

where $E_2(\tau_2(q, \eta_2; \delta_1, x, A_1(q_s))) = \int \tau_2(q, \eta_2; \delta_1, x, A_1(q_s)) d_{\Phi(\eta_2)} = \tau_2(q|\delta_1, x, A_1(q_s))$. The optimal advertising spending for period 1, $a_1^* = A_1(q_s; x)$, satisfies the equilibrium condition

$$\frac{\partial [\pi_1(a_1; S_1) + E_1 V(S_2|a_1, S_1)]}{\partial a_1}|_{a_1 = a_1^*} = 0 \tag{8}$$

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, I try to keep the model trackable and still be able to investigate the questions in interest. Second, the movie's box office performance positively impacts the revenue from other nontheatrical windows. Third, distributors and exhibitors normally have a long-term relationship for many movies and they have many 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.

4.5 Advertising-Watching Equilibrium

Since this dissertation 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) $A_1(\cdot) \in \arg\max[\pi_1(a_1; S_1) + E_1 V(S_2|a_1, S_1)]$ *and* $A_2(\cdot) \in \arg\max[\pi_2(a_2; S_2)]$

(E2) $w_{it}(\cdot) = 1$ *if and only if* $\gamma x + \alpha TD + E^c_{it}[q|I_i(t)] - p + \eta_t + \varepsilon_{it} \geq 0$.

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

Lemma 2 *If the advertising policy function $A_1(q_s, x)$ is increasing in q_s, then $Eu_{it} = \gamma x + \alpha TD + E^c_{i1}[q \mid k_{i1}, A_1(q_s)] - p + \eta_1 + \varepsilon_{it}$ is increasing in k_{i1}, and the best response rule is*

$$w_{i1}(k_{i1}; A_1(s)) = \begin{cases} 1 & if \quad k_{i1} \geq \lceil k^*_{i1} \rceil \\ 0 & if \quad k_{i1} < \lceil k^*_{i1} \rceil \end{cases}$$

where k^*_{i1} is defined by $\gamma x + \alpha TD + E^c_{i1}[q \mid k_{i1}, A_1(q_s)] - p + \eta_1 + \varepsilon_{it} = 0$, and $\lceil k^*_{i1} \rceil$ is the smallest integral which is not smaller than k^*_{i1}. Therefore $\tau_1(a_1, \eta_1; x, A_1(q_s))$ is an increasing function of a_1 conditional on x.

Proof. Since $\widetilde{k}_{i1} \sim pois(\lambda a_1)$ and the family of Poisson distributions satisfy MLRP, the posterior CDF of a_1, $H_{i1}(a_1|k_{i1})$, is a decreasing function of k_{i1}. Since $A_1(\cdot)$ is an increasing smooth function of q_s, the posterior CDF of q, $G_{i1}(q|k_{i1})$, is also a decreasing function in k_{i1}. If $k_{i1} > k_{h1}$, then $\Pr[q > \alpha|k_{i1}] > \Pr[q > \alpha|k_{h1}]$, so $q|k_{i1}$ first-order stochastically dominates $q|k_{h1}$ and $E[q|k_{i1}] > E[q|k_{h1}]$. Therefore Eu_{i1} increases in k_{i1} and $\tau_1(a_1, \eta_1; x, A_1(q_s))$ is an increasing function of a_1. ■

Lemma 3 *$Eu_{i2} = \gamma x + \alpha TD + E^c_{i2}\left[q \mid \overline{q}_{iw}, \delta_1, A_1(q_s)\right] - p + \eta_2 + \varepsilon_{i2}$ is increasing in $\overline{q}_{iw}$ and $\tau_2(q, \eta_2; \delta_1, x, A_1(q_s))$ is increasing in q.*

Proof. This only requires $E^c_{i2}\left[q \mid \overline{q}_{iw}, \delta_1, A_1(q_s)\right]$ increases in $\overline{q}_{iw}$. The assumption that $\overline{q}_{iw} \sim iidN\left(q, \frac{\sigma_w^2}{\rho\delta_1}\right)$ ensures this lemma holds.[10] ■

Assumption A1: $\frac{\delta_1}{M}$ is small.

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 very important for a movie's success. This should be a very reasonable assumption since it is supported by data. Table 4 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 4, 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

[10] As long as the family of distributions $f\left(\overline{q}_{iw} \mid q, \frac{\sigma_w^2}{\rho\delta_1}\right)$ has MLRP in q, then this lemma holds.

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: $\frac{\tau_2(q,\eta_2;\delta_1,x,A_1(q_s))}{\partial\delta_1\partial q}$ is nonnegative or limited negative.

In period 2, consumers have two information sources from which to update their beliefs: the movie's market share in period 1 and the WOM among consumers about the movie. δ_1 is determined by the studio's advertising action and, therefore, can be called as "firm-generated" information. $\overline{q}_{iw}$ 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 substitutes, the ratio is small enough.

Lemma 4 *If $A_1(\cdot)$ is a best response to $w_{it}(\cdot)$, then with assumptions A1 and A2, $A_1(\cdot)$ is nondecreasing, and, for $q_s \in Q_S^{'} \subseteq Q_S$[11], $A_1(\cdot)$ 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

[11] Q_s is the domain of random variable q_s and $Q_s^{'}$ is the subset of the domain of q_s.

rewrite the studio's revenue maximization problem in the following way:

$$\max_{a_1}\Pi_0\left(a_1, q_s\right) = \int \pi\left(a_1, q\right) g\left(q \mid q_s\right) dq$$

$$\begin{aligned}
s.t.\pi\left(a_1, q\right) &= \int \left[\delta_1 p + \left(\varphi_2\left(a_1, a_2\right) M - \delta_1\right) p \int \tau_2\left(q, \eta_2, \delta_1\right) d_{\Phi(\eta_2)}\right] d_{\Phi(\eta_1)} \\
&\quad -a_1 - a_2 \\
\delta_1 &= \Delta_1\left(a_1, \eta_1\right) = \varphi_1\left(a_1\right) M \tau_1\left(a_1, \eta_1\right) \\
a_2 &= \max\left(\frac{\ln\left(\lambda M p E_2\left(\tau_2\left(q, \eta_2, \delta_1\right)\right)\right)}{\lambda} - \mu a_1, 0\right) = A_2\left(a_1\right) \\
\Pi_0 &\geq 0 \\
a_1 &\geq 0 \\
a_2 &\geq 0
\end{aligned}$$

Theorem 2 in Athey (2002) requires that $\pi\left(a_1, q\right)$ satisfies SC2 in $\left(a_1, q\right)$ and $g\left(q \mid q_s\right)$ is log-spm[12] as a minimal pair of sufficient conditions for $A_1\left(q_s\right)$ to be nondecreasing in q_s. $g\left(q \mid q_s\right)$ is log-spm can be met by the assumption that $g\left(q \mid q_s\right)$ is conditional normal distribution and has MLRP in q_s. By assuming $\pi\left(a_1, q\right)$ is C^2, I just need to check the sign of $\frac{\partial^2 \pi(a_1,q)}{\partial a_1 \partial q}$.

$$\begin{aligned}
\frac{\partial^2 \pi\left(a_1, q\right)}{\partial a_1 \partial q} &= \left(\frac{\varphi_2\left(a_1, a_2\right)}{a_1} M - \frac{\partial \delta_1}{\partial a_1}\right) p \int \frac{\partial \tau_2\left(q, \eta_2, \delta_1\right)}{\partial q} d_{\Phi(\eta_2)} \\
&\quad + \left(\varphi_2\left(a_1, a_2\right) M - \delta_1\right) p \int \frac{\partial^2 \tau_2\left(q, \eta_2, \delta_1\right)}{\partial \delta_1 \partial q} \frac{\partial \delta_1}{\partial a_1} d_{\Phi(\eta_2)}
\end{aligned}$$

With assumptions A1 and A2, we know $\frac{\partial^2 \pi(a_1,q)}{\partial a_1 \partial q} \geq 0$. Then by using Theorem 2 in Athey (2002), we know $A_1\left(\cdot\right)$ is nondecreasing. With assumption A2 and the

[12]log-supermodular is abbreviated to log-spm. Here, it means that q_s shifts the conditional distribution of q according to the monotone likelihood ratio property.

assumption that $f(q_s \mid q)$ satisfies MLRP, $A_1(\cdot)$ cannot be constant for all $q_s \in Q_S$, then $A_1(\cdot)$ is increasing for $q_s \in Q_S' \subseteq Q_S$. I will briefly discuss assumption A2 and the conditions under which $\frac{\partial^2 \pi(a_1, q)}{\partial a_1 \partial q} \geq 0$ is satisfied in Appendix 8 ■

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

(E3) $w_{i1}(k_{i1}; A_1(s)) = \begin{cases} 1 & if \quad k_{i1} \geq \lceil k_{i1}^* \rceil \\ 0 & if \quad k_{i1} < \lceil k_{i1}^* \rceil \end{cases}$ *where* k_{i1}^* *is defined by* $\gamma x + \alpha TD + E_{i1}^c [q \mid k_{i1}, A_1(q_s)] - p + \eta_1 + \varepsilon_{it} = 0$, *and* $\lceil k_{i1}^* \rceil$ *is the smallest integral which is not smaller than* k_{i1}^*. *Then* $\tau_1(a_1, \eta_1; x, A_1(q_s))$ *is an increasing function of* a_1.

(E4) $A_1(q_s)$ *is nondecreasing in* $q_s \in S$, *and for* $q_s \in S' \subseteq S$, $A_1(\cdot)$ *is increasing.*

4.6 Simplified Examples

4.6.1 Discrete Type Example

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 μ_0 to be q_H and $(1 - \mu_0)$ to be q_L. The studio observes the quality signal q_{sj} (either q_{sH} or q_{sL}), and the probability to be right is $\eta \in \left(\frac{1}{2}, 1\right)$, which means $\Pr(q_{sH}|q_H) = \Pr(q_{sL}|q_L) = \eta$ and $\Pr(q_{sH}|q_L) = \Pr(q_{sL}|q_H) = 1 - \eta$. After receiving q_{sj}, the studio updates its belief of quality and decides the advertising spending a_{j1}. First, I assume that when the studio receives q_{sH}, it will spend A_{H1}; when it receives q_{sL}, 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 $\widetilde{k}_{ij1} \sim pois\left(A_{j1}\right)$ ($j = H$ or L). The probability that $q_j = q_H$ perceived by consumer i is updated to be

$$\mu_1\left(k_{ij1}\right) = prob\left(q_H|k_{ij1}\right) = \frac{prob\left(k_{ij1}|q_H\right)prob\left(q_H\right)}{prob\left(k_{ij1}|q_H\right)prob\left(q_H\right) + prob\left(k_{ij1}|q_L\right)prob\left(q_L\right)}$$

where $prob\left(k_{ij1}|q_H\right) = prob\left(k_{ij1}|A_{H1}\right)prob\left(A_{H1}|q_H\right) + prob\left(k_{ij1}|A_{L1}\right)prob\left(A_{L1}|q_H\right) = prob\left(k_{ij1}|A_{H1}\right)\eta + prob\left(k_{ij1}|A_{L1}\right)(1-\eta)$ and $prob\left(k_{ij1}|q_L\right) = prob\left(k_{ij1}|A_{H1}\right)(1-\eta) + prob\left(k_{ij1}|A_{L1}\right)\eta$. And it is easy to show that $\mu_1\left(k_{ij1}\right)$ is an increasing function of k_{ij1}. So consumer i's expected utility[13]

$$Eu_{ij1} = E_{ij1}\left[q_j \mid I_i(1)\right] - p = \left(\mu_1\left(k_{ij1}\right)q_H + \left(1-\mu_1\left(k_{ij1}\right)\right)q_L\right) - p \geq 0$$

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 $\rho\delta_{j1}$ of consumers will know the true quality of the movie, and $(1-\rho\delta_{j1})$ fraction of consumers will keep their prior perceived quality level $\overline{q} = \mu_0\, q_H + (1-\mu_0)\, q_L$. WOM communication ratio, ρ, is used to indicate the fraction of consumers who like to share their review of the movie with other consumers after watching it, and δ_{j1} is the

[13]Here, consumers' expected utility only depends on expected quality and price for simplicity's sake.

market share in period 1.[14] Consumer i's utility in period 2 is

$$Eu_{ij2} = E_{ij2}\left[q_j \mid I_i(2)\right] - p = \begin{array}{l} q_j - p, \text{ if she learns from WOM, } j = H \text{ or } L \\ \overline{q} - p, \text{ otherwise} \end{array}$$

Here, I assume $\overline{q} - p < 0$ and $q_H - p > 0$. Then the studio with q_j has market share in period 1 $\delta_{j1} = prob\left(k_{ij1} > 0\right) porb\left(k_{ij1} \geq k_1^* | A_{j1}\right)$. In period 2, the studio with q_H has market share $\delta_{j2} = prob\left(k_{ij2} > 0\right) \rho * \delta_{j1}$ 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,

$$E_0\left[\delta_{H2} | q_{sH}\right] = prob\left(q_H | q_{sH}\right) * \delta_{H2}$$

and

$$E_0\left[\delta_{L2} | q_{sL}\right] = prob\left(q_H | q_{sL}\right) * \delta_{H2}$$

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_{sH} is:

$$\underset{A_{H1},A_{H2}}{Max}\ \Pi_H = \delta_{H1}\left(A_{H1}\right) N_1 p + E_0\left[\delta_{H2}\left(A_{H1}, A_{H2}\right) | q_{sH}\right] N_2 p - A_{H1} - A_{H2}$$

[14] 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.

The profit maximization problem for the studio receiving q_{sL} is:

$$\underset{A_{L1},A_{L2}}{Max}\,\Pi_L = \delta_{L1}\left(A_{L1}\right)N_1p + E_0\left[\delta_{L2}\left(A_{L1},A_{L2}\right)|q_{sL}\right]N_2p - A_{L1} - A_{L2}$$

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 $A^*_{H1} > A^C_{H1}$ and $A^*_{L1} > A^C_{L1} = 0$; 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 cases, there is no market.)

4.6.2 Continous Type Example

To gain more intuition about the role of information uncertainty in determining studios' advertising spending and how studios' equilibrium advertising strategies change if the WOM communication level changes, I simplify the general model in the following ways. First, if consumers in period 1 do not infer a movie's quality from observed advertising intensity, then $\tau_1(\cdot)$ is independent of a_1 and a_1 is only used to reach consumers. Second, the distributor of a new movie only decides a_1 instead of $(a_1,\ a_2)$ to reach consumers. Therefore, the distributor's profit maximization problem becomes a one-period problem instead of a two period problem. Those two simplifications help simplify the computation of both consumers' and distributors' optimization problems and attain numerical results. Then, with different assumptions about information asymmetry between consumers and distributors, I discuss the following three cases:

Case 1: There is no information asymmetry in the pre-release period, so consumers know q_s as well as the distributor and $E_{it}[q \mid I_i(1)] = E_{it}[q \mid q_s]$. In the post-release period, consumers learn about true q from WOM information $\overline{q}_{iw}$. Since they already know q_s, there is no information learning from market share of release week (δ_1). When deciding a_1, the distributor considers the impact of δ_1 on $\frac{\sigma_w^2}{\rho\delta_1}$ (higher δ_1 leads to smaller variance and more accurate WOM information).

Case 2: There is information asymmetry between consumers and the distributor in pre-release period, so $E_{it}[q \mid I_i(1)] = \overline{q}$. In the post-release period, consumers not only learn about true q from WOM information $\overline{q}_{iw}$, but also learn about q_s from market share δ_1 ($\delta_1 = \delta_1(A_1(q_s))$ is an increasing function in q_s). When deciding a_1, the distributor only considers the impact of δ_1 on $\frac{\sigma_w^2}{\rho\delta_1}$ (higher δ_1 leads to smaller variance and more accurate WOM information), but ignore that δ_1 can be a signal of q_s in the post-release period in a separating equilibrium.

Case 3: This case is similar to Case 2, except that the distributor considers both the impact of δ_1 on $\frac{\sigma_w^2}{\rho\delta_1}$ and the signaling effect of δ_1 when deciding the optimal advertising a_1.

Figure 6-1 shows us the numerical results of the equilibrium advertising strategy $a_1 = A_1(q_s)$ for those three cases. By comparing those three curves, we can observe the following three interesting results. First, all three curves are increasing (nondecreasing for case 1) in q_s, which shows the existence of a separating equilibrium. In particular, the green curve for case 3 shows the existence of a signaling equilibrium for $\delta_1 = \delta_1(A_1(q_s))$. Second, the blue and green curves (case 2 and 3) show us that movies with very low q_s can enter the market because of information asymmetry between consumers and the distributor in the pre-release period. Meanwhile, the black curve (case 1) shows us that movies whose q_s is sufficiently low (below some threshold value which is around 4.5 in the example here) do not enter the market when there is no information asymmetry between consumers and the distributor. Therefore, the comparison of those cases helps us understand the important impact of information structure on the market structure: if the information asymmetry between consumers and the distributor can be reduced by the pre-release signaling effect of advertising, fewer low-quality movies will be provided to consumers in the market. Third, the blue and green curves' shapes are very similar, but the green curve is always above the blue curve, which shows the cost of having a signaling equilibrium. Once distributors understand that consumers can infer information about q_s from δ_1 in the post-release period, then distributors with low-quality movies will spend more to pretend their movies are high quality, and distributors with high-quality movies are pushed to spend more to differentiate their movies from low-quality movies. This process helps to form the "inflation" of advertising spending for all levels of q_s.[15]

[15] In this simplified example, we only discuss the signaling role of $\delta_1 = \delta_1(A_1(q_s))$ in the post-

But how does the distributor's equilibrium advertising strategy change with higher levels of WOM communication? Figure 6-2 shows us the change of $a_1 = A_1(q_s)$ when I set higher value for ρ. The shapes of those curves are the same as in Figure 6-1, which shows us the similar results discussed above. The main difference is the vertical gap between the green curve and the blue curve: with higher value of ρ, the signaling equilibrium requires lower advertising spending for all levels of q_s. Therefore, there is less advertising spending "inflation" required for signaling purposes with better WOM communication among consumers in the post-release period. The reason is very intuitive: in the post-release period, consumers have two information sources from which to learn about a movie's quality — $\overline{q}_{iw}$ and δ_1. When $\overline{q}_{iw}$ becomes more accurate (lower variance), consumers will put more weight on WOM instead of δ_1. Then there is less motivation for distributors to signal through δ_1.[16]

Figure 6-3 and Figure 6-4 show the distributor's optimized profit as a function of q_s with low and high values of ρ. Movies with higher quality ($q_s > 6$) earn higher profits in the information symmetry case than in the information asymmetry case (the black curve is above the blue and green curves); the opposite is true for movies of lower quality ($q_s < 6$), (black curve is below the blue and green curves). The vertical gap between blue curve and the green curve also shows the cost of having a signaling equilibrium. Higher level of ρ helps to shrink the profit gap, therefore distributors can actually make more money when there is better WOM communication among consumers.

release period, however, if we consider the signaling role of $a_1 = A_1(q_s)$ in the pre-release period, there should be more inflation of advertising spending in the signaling equilibrium.

[16] Again, we only discuss the signaling role of $\delta_1 = \delta_1(A_1(q_s))$ in the post-release period in this simplified example. If we consider the signaling role of $a_1 = A_1(q_s)$ in the pre-release period, the change of WOM communication level should reduce the inflation as well, but the reason is different. Actually consumer learning about q through WOM is the main mechanism which helps to form the signaling equilibrium of $a_1 = A_1(q_s)$. Therefore, higher level of WOM communication will enhance the signaling role of advertising spending and require lower level of advertising spending for all q_s levels.

5 Estimation Strategy

In this dissertation, I consider estimating the structural parameters of the proposed model using the method of maximum-likelihood estimation. First, I derive the logarithm of the likelihood function and formulate the maximum likelihood estimation problem. Then, I describe the MPEC algorithm used for implementing the ML estimator. The identification of parameters is briefly discussed at the end of this chapter.

5.1 Likelihood Contribution

From data, I observe box office performance and advertising spending in two periods for movie j: $Y_j = (\delta_{j1}, \delta_{j2}, a_{j1}, a_{j2})'$. From equations (5), (6), (7) and (8), we know that the observed variables Y_j can be expressed as functions of unobserved random variables $Z_j = \left(\eta_{j1}, \eta_{j2}, q_{js}, q_j\right)'$ in a more compact form. The relationship between Y_j and Z_j can be devoted to be $Y_j = \zeta\left(Z_j | x_j, \Theta\right)$ which is also the Bayesian-Nash equilibrium equation. So unobserved random variables can be written as $Z_j = \zeta^{-1}\left(Y_j | x_j, \Theta\right)$, where $\Theta = \left\{\left\{\gamma, \alpha, \sigma^2_{\eta_1}, \sigma^2_{\eta_2}\right\}, \left\{\overline{q}, \sigma^2_q, \sigma^2_s, \frac{\sigma^2_w}{\rho}\right\}, \{\lambda, \mu\}\right\}$ is denoted as the set of structural parameters.

If we assume that $Z_j = (\eta_1, \eta_2, q_{js}, q_j)' \sim MVN\left(U, \Sigma\right)$, where MVN stands for multivariate normal, then the pdf of Z_j is given by

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

By using the standard transformation of variables technique, we obtain the joint density of $Y_j = (\delta_{j1}, \delta_{j2}, a_{j1}, a_{j2})$ for the movie j as follows:

$$\begin{aligned} m\left(Y_j|x_j,\Theta\right) &= g_z\left(\zeta^{-1}\left(Y_j|x_j,\Theta\right)\right)\left|J_{(Z\to Y)}\right| \\ &= \frac{1}{(2\pi)^2}\left|\Sigma\right|^{-\frac{1}{2}}\exp\left(-\frac{1}{2}\zeta^{-1}\left(Y_j|x_j,\Theta\right)'\Sigma^{-1}\zeta^{-1}\left(Y_j|x_j,\Theta\right)\right) \\ &\quad \left\|\frac{\partial\zeta^{-1}\left(Y_j|x_j,\Theta\right)}{\partial Y_j}\right\| \end{aligned} \tag{9}$$

where $J_{(Z\to Y)}$ is the Jabobian matrix which is derived in the Appendix 8. The joint likelihood function is written as:

$$\begin{aligned} L &= \prod_{j=1}^{J} m\left(Y_j|x_j,\Theta\right) \\ &= (2\pi)^{-2J}\left|\Sigma\right|^{-\frac{J}{2}}\exp\left(-\frac{1}{2}\sum_{j=1}^{J}\zeta^{-1}\left(Y_j|x_j,\Theta\right)'\Sigma^{-1}\zeta^{-1}\left(Y_j|x_j,\Theta\right)\right) \\ &\quad \prod_{j=1}^{J}\left\|\frac{\partial\zeta^{-1}\left(Y_j|x_j,\Theta\right)}{\partial Y_j}\right\| \end{aligned}$$

Then the log-likelihood is

$$\begin{aligned} \ln L &= -2J\ln(2\pi) - \frac{J}{2}\ln\left|\Sigma\right| - \frac{1}{2}\sum_{j=1}^{J}\zeta^{-1}\left(Y_j|x_j,\Theta\right)'\Sigma^{-1}\zeta^{-1}\left(Y_j|x_j,\Theta\right) \\ &\quad +\sum_{j=1}^{J}\ln\left\|\frac{\partial\zeta^{-1}\left(Y_j|x_j,\Theta\right)}{\partial Y_j}\right\| \end{aligned}$$

Solving $\frac{\partial \ln L}{\partial \Sigma} = 0$ for Σ, we get

$$\Sigma = \frac{1}{J}\sum_{j=1}^{J}\zeta^{-1}\left(Y_j|x_j,\Theta\right)\zeta^{-1}\left(Y_j|x_j,\Theta\right)'$$

Then the concentrated log-likelihood function is

$$\widehat{\ln L} = \sum_{j=1}^{J}\ln\left\|\frac{\partial\zeta^{-1}\left(Y_j|x_j,\Theta\right)}{\partial Y_j}\right\| - \frac{J}{2}\ln\left|\frac{1}{J}\sum_{j=1}^{J}\zeta^{-1}\left(Y_j|x_j,\Theta\right)\zeta^{-1}\left(Y_j|x_j,\Theta\right)'\right| \tag{10}$$

The maximum likelihood estimation problem is formulated as

$$\max_{\Theta} \widehat{\ln L} \left(Y, x; \Theta, \zeta\left(\Theta\right)\right) \tag{11}$$

and ML estimator is defined as

$$\Theta^{MLE} = \arg\max_{\Theta} \left\{ \max_{\zeta(\Theta)} \widehat{\ln L} \left(Y, x; \Theta, \zeta\left(\Theta\right)\right) \right\}$$

5.2 Estimation Method

To evaluate the likelihood function, I have to solve the advertising policy function $a_1^* = A\left(q_s; x\right)$ as the equilibrium result of the incomplete information game between studios and consumers. In this dissertation, I assume movie studios of new movies make their advertising decisions in the context of competitive monopoly 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 (11). The general idea about implementing the NFXP algorithm is that it involves two loops: in the outer-loop, search the structural parameter space over Θ to maximize $\left\{ \max_{\zeta(\Theta)} \widehat{\ln L} \left(Y, x; \Theta, \zeta\left(\Theta\right)\right) \right\}$; in the inner-loop, for any given values of Θ, 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 Θ, 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 Θ, 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 dissertation 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 dissertation, 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 11 can be reformulated as a constrained optimization problem in the joint space of structural parameters and economic equilibrium as the following:

$$\begin{aligned} \max_{\Theta, Z, \zeta(\cdot)} \quad & \widehat{\ln L} \left(Y, x; \Theta, \zeta \left(\Theta \right) \right) \\ s.t. \quad & Y = \zeta \left(Z | x, \Theta \right) \end{aligned} \tag{12}$$

where equilibrium equations $Y = \zeta(Z|x, \Theta)$ are written as constraints, and structural parameters Θ, 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 $Z_j = \zeta^{-1}(Y_j|x_j, \Theta)$ for each observation and each guess of Θ, 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 Equations (equation (7)) have closed-form expressions, but the pre-release advertising policy function $a_1^* = A(q_s; x, \Theta)$ cannot be written in an analytical format explicitly. However, the first-order condition with respect to a_1 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 Θ. The first-order condition used as an equilibrium condition plays the same role as the Bellman equation in dynamic games. 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_s, conditional on each observed value of x, which brings too much computational burden. Therefore, instead of considering all possible values, the studios' first-order conditions are solved at a subset of points in the support of q_s and the policy function is approximated using Chebyshev polynomials[17].

In my model, consumers are assumed to be fully rational so that they understand the signaling mechanism and infer q_s from observed advertising intensity and market share. This means the equilibrium advertising policy function influence consumers' watching decisions through their utility function. Therefore, instead of approximating advertising policy function, I approximate the inverse function, $q_s = H(a_1; x)$, of $a_1^* = A(q_s; x)$. Further, all observed characteristics of a new movie included in x 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 $X = \gamma x$, 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 Chebysheve polynomials for a_1 and nX be the order of Chebysheve polynomials for X. The inverse advertising function $q_s = H(a_1; X) : [\underline{a_1}, \overline{a_1}] \otimes [\underline{X}, \overline{X}] \rightarrow R$ is approximated by $\widehat{H}(a_1; X) = \kappa' \Lambda(a_1; X)$, where κ is $N * 1$ vector of approximation parameters, $\Lambda(a_1; X)$ is $N * 1$ vector of N Chebyshev polynomials and $N = (na + 1)(nX + 1)$. a_m ,x_m are grids of $ma \geq na + 1$

[17]Chebyshev polynomials are used for the approximation to maximize the stability of the approximation to the policy functions and avoid Runge's oscillatory phenomenon.

and $mX \geq nX + 1$ Chebyshev nodes on $\left[\underline{a_1}, \overline{a_1}\right] \otimes \left[\underline{X}, \overline{X}\right]$.

With $q_s = \widehat{H}(a_1; X)$, the equilibrium equation (8) can be approximated by

$$\frac{\partial \Pi\left(a_1, \widehat{H}(a_1; X); x, \Theta\right)}{\partial a_1} + \frac{\partial \Pi\left(a_1, \widehat{H}(a_1; X); x, \Theta\right)}{\partial \widehat{H}(a_1; X)} \frac{\partial \widehat{H}(a_1; X)}{\partial a_1} \approx 0 \tag{13}$$

where $\Pi\left(a_1, \widehat{H}(a_1; X); x, \Theta\right) = \pi_1(a_1; S_1) + E_1 V(S_2 | a_1, S_1)$. Note that the pre-release advertising a_1 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

$$\begin{aligned}
& \max_{\Theta, \{Z_j\}_{j=1}^{J}, \kappa} \widehat{\ln L}(Y, x; \Theta, \zeta(\Theta)) \\
s.t.\ q_{js} &= \widehat{H}(a_{j1}, X_j) = \kappa' \Lambda(a_{j1}, X_j) \\
\delta_{j1} &= \Delta_1\left(a_{j1}, \eta_{j1}; x_j, \Theta, \widehat{H}(a_{j1}, X_j)\right) \\
a_{j2} &= A_2\left(q_j; a_{j1}, \delta_{j1}, x_j, \Theta, \widehat{H}(a_{j1}, X_j)\right) \\
\delta_{j2} &= \Delta_2\left(a_{j2}, q_j, \eta_{j2}; a_{j1}, \delta_{j1}, x_j, \Theta, \widehat{H}(a_{j1}, X_j)\right) \quad for\ j = 1, 2...., J \\
0 &= \frac{\partial \Pi\left(a_{m1}, \widehat{H}(a_{m1}; X_m); X_m, \Theta\right)}{\partial a_{m1}} + \frac{\partial \Pi\left(a_{m1}, \widehat{H}(a_{m1}; X_m); X_m, \Theta\right)}{\partial \widehat{H}(a_{m1}; X_m)} \\
& \frac{\partial \widehat{H}(a_{m1}; X_m)}{\partial a_{m1}} \\
for\quad m &= 1, 2, ..., (ma * mX)
\end{aligned}$$

Here, the structural parameters Θ, $\left\{Z_j = \left(\eta_{j1}, \eta_{j2}, q_{js}, q_j\right)'\right\}_{j=1}^{J}$ and approximation parameters κ are chosen to maximize the likelihood function. Integrals over demand shocks in period 1, η_1, and true unobserved quality, q, are approximated by using 20 draws from their distributions with antithetic acceleration. Integrals over demand shocks in period 2, η_2, and WOM signal $\overline{q}_{iw}$ are approximated by using Gauss–Hermite quadrature[18] 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_1 and 3 degree of Chebyshev polynomials for X. Then the total number of approximation parameters to be estimated is 20. Equilibrium condition (13) is evaluated at $ma = 5$ points in the domain of $[0.68, 42]$ for a_1 and $mX = 4$ points in the domain of $[0, 2.5]$ for X.[19] To increase the computation accuracy and reduce the estimation time, I provided the hand-coded 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 $\Theta = \left\{\left\{\gamma, \alpha, \sigma^2_{\eta_1}, \sigma^2_{\eta_2}\right\}, \left\{\overline{q}, \sigma^2_q, \sigma^2_s, \frac{\sigma^2_w}{\rho}\right\}, \{\lambda, \mu\}\right\}$. There are three types of parameters: demand preference parameters $\left\{\gamma, \alpha, \sigma^2_{\eta_1}, \sigma^2_{\eta_2}\right\}$; information structure parameters $\left\{\overline{q}, \sigma^2_q, \sigma^2_s, \frac{\sigma^2_w}{\rho}\right\}$, advertising parameters (supply side parameters) $\{\lambda, \mu\}$. I will discuss their identification in turn.

Because I can only observe data from one market (the U.S. domestic market) over time for each movie, I need to assume preference parameters for observed characteristics, γ, are the same for every consumer. As mentioned before, I assume there are only two periods: opening week and post-release weeks to make estimation easier. For post-release weeks, I aggregate the box office performance and advertising

[18] Quadratures are used instead of simulation because they performs much better when compared with the results of simulation, and allowed for much faster execution.

[19] For the domain of a_1, I use the observed range of advertising spending in period 1 in the data. For the domain of X, I first try a large enough range and then adjust the range to appropriate values to improve the estimation accuracy.

spending data together and use a time dummy variable to capture the demand's level difference between period 1 and period 2. The variance in advertising spending a_1 and a_2 corresponding to the variance in x can be used to identify γ. Conditional on a_1, a_2 and x, the level difference between δ_1 and δ_2 across all movies can be used to identify α. For aggregate demand shocks, the variance in δ_1 conditional on x and a_1, can be used to identify $\sigma^2_{\eta_1}$, and the variance of δ_2 conditional on δ_1, a_1, x, and a_2 can be used to identify $\sigma^2_{\eta_2}$. The distribution parameters,$\left(\overline{q}, \sigma^2_q\right)$, for movie's unobserved quality, q, can identified by mean and variance of a_2 conditional on x, a_1 and δ_1. The noisy signal variance σ^2_s can be identified by the variance of a_1 conditional on a_2 after $\left(\overline{q}, \sigma^2_q\right)$ become known. The adjusted WOM variance parameter $\frac{\sigma^2_w}{\rho}$ can be identified by covariance of $\delta_1$and δ_2 conditional on a_1 and a_2 . Note that the information transmission speed parameter, ρ, cannot be separately identified from WOM variance σ^2_w, but assuming σ^2_w as a constant over time is a reasonable assumption. "Reaching efficiency" parameter λ in the adverting reach function $\varphi_t\left(\cdot\right)$ can be identified by covariance of δ_2 and a_2 conditional on $\tau_2\left(\cdot\right)$. Then advertising "depreciation parameter" μ can be identified by covariance of δ_2 and a_1 conditional on $\tau_2\left(\cdot\right)$ when λ becomes known.

6 Empirical Results

6.1 Estimates

Advertising's Signaling Effect The estimated inverse advertising policy function is presented in Figure 7 and Table 5. When I do the constrained MLE estimation, I don't impose any shape constraints on advertising function, but only require that the first order conditions of studios' profit optimization problem to be satisfied. Figure 7 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 5 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, if we check the dataset, only the scenarios shown in the lower right corner of Table 5 happened in the real world. Movies with high value of X usually have high value of q_s as well, and therefore the advertising spending is still an increasing function of q_s.

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 that correlation between a_1 and q_s. For the model estimated, I assume consumers are fully rational so that they understand the signaling mechanism and infer unobservable product quality from the advertising. On the other hand, I can also assume consumers are limited rational, which means that they only learn about product existence and observable attributes through advertising. In this case, equation (1) becomes

$$E_i^c \left[q \mid I(1) \right] = \overline{q}$$

and advertising is only used to reach consumers.

I do the estimation with limited rationality assumption about consumers and get the corresponding maximized likelihood. In a comparison of likelihood values, the original model is preferred, which supports the existence of advertising's signaling effect[20].

Utility Function Parameters The estimates of the preference parameter in consumer utility function is reported Table 6.1. 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 different types of distributors attract consumers in different ways. 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

[20] Here I measure the relative quality of models for a given set of data by using Akaike Information Criterion (AIC).

the quality of movies. To some extent, my results are consistent with his arguments.

Information Learning Parameters Table 6.2 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, σ_s^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 σ_s^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 $\beta_q^s(0) = \frac{\sigma_q^2(0)}{\sigma_q^2(0)+\sigma_s^2} = \frac{\sigma_q^2}{\sigma_q^2+\sigma_s^2} = 0.09$, and the updated variance becomes $\sigma_q^{s2}(1) = \frac{\sigma_q^2\sigma_s^2}{\sigma_q^2+\sigma_s^2} = 0.309$. 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 pass information about a movie's true quality. On average[21], consumers put around $\beta_q^c(1) = \frac{\sigma_q^{c2}(1)}{\sigma_q^{c2}(1)+\frac{\sigma_w^2}{\rho\delta_1}} = \frac{\sigma_q^{s2}(1)}{\sigma_q^{s2}(1)+\frac{\sigma_w^2}{\rho\delta_1}} = 0.93$ 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 $\sigma_q^{c2}(2) = \frac{1}{\frac{1}{\sigma_q^{c2}(1)}+\frac{\rho\delta_1}{\sigma_w^2}} = \frac{1}{\frac{1}{\sigma_q^{s2}(1)}+\frac{\rho\delta_1}{\sigma_w^2}} = 0.02$ 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 6.3. With λ equaling to 0.131, the market

[21] Mean(δ_1) is normalized to 1.

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 μ shows that it is easy for consumers to forget about the movie when they enter period 2. 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 new 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:

$$X^2 = \sum_{k=1}^{K} \frac{(n_k^o - n_k^e)^2}{n_k^e} \sim X_{K-1}^2$$

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. $n_k^e = p_k N$ 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 7-1 shows the observed and expected numbers of data fall in each cells for both 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 adverting 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 $a1$ and X separately and form cross-product cells, where $a1$ 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 7-2 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 structure.

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 learning 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_s before the release of its new movie, 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

$$E_{i1}^{c}\left[q \mid k_{i1}, A_1\left(q_s\right)\right] = \int q \cdot g_{i1}\left(q|k_{i1}, A_1\left(q_s\right)\right) d_q$$

for period 1 and

$$E_{i2}^{c}\left(q\right) = E_{i1}^{c}\left[q \mid \delta_1\right] + \beta_q^c\left(1\right)\left(\overline{q}_{iw} - E_{i1}^{c}\left[q \mid \delta_1\right]\right)$$

for period 2. Both of them are affected by advertising spending a_1 through its signaling effect. The demand in period 1, $\delta_1 = \varphi_1\left(a_1\right) M\tau_1\left(a_1;\cdot\right)$, shows that a_1 affects demand both through reaching channel $\left(\varphi_1\left(a_1\right)\right)$ and signaling channel $\left(\tau_1\left(a_1;\cdot\right)\right)$.

For counterfactual experiment 1, I assume consumers automatically know q_s as well as the studio after the movie is released. Therefore, consumer i's perceived expected quality of movie j becomes

$$E_{i1}^{c}\left(q\right) = E_{1}^{s}\left(q\right) = \overline{q} + \beta_q\left(0\right)\left(q_s - \overline{q}\right)$$

in period 1 and

$$E_{i2}^{c}\left(q\right) = E_{i1}^{c}\left(q\right) + \beta_q^c\left(1\right)\left(\overline{q}_{iw} - E_{i1}^{c}\left(q\right)\right)$$

in period 2. Both are independent of studios' advertising spending and the equilibrium advertising strategies. The demand in period 1 becomes $\delta_1 = \varphi_1(a_1) M \tau_1(\cdot)$ which a_1 affects only through reaching channel ($\varphi_1(a_1)$).

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_s 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

$$E_{i1}^c(q) = \overline{q}$$

in period 1 and

$$E_{i2}^c(q) = \overline{q} + \beta_q^c(1)(\overline{q}_{iw} - \overline{q})$$

in period 2. Similar to experiment 1, a_1 affects the demand only through its reaching effect.

The results of these two experiments are reported in Table 8 and Figure 8. 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 8, the advertising spending pattern is very similar for experiment 1 and experiment 2. This is because we check the average advertising spending across all movies in my sample. For movies with different characteristics (observed and unobserved), studios' advertising strategies are very different under those two different information structures. In Figure 8, 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 9 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 roles of advertising and WOM communication among consumers have not been fully explored. This research structurally models movie studios' optimal advertising strategies and consumer information learning in an equilibrium setting in the motion picture market 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 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 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 dissertation, 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 dissertation, 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 alternative 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.

Appendix

A.1: Assumption 2

In this section, I will briefly discuss some of the conditions under which A2 is a reasonable assumption, as well as how A2 can be used to support the proof of Lemma 4. Several model simplifications are made without loss of generality to provide more intuition.

Consumer i's expected utility from watching the movie in period t $(t = 1, 2)$ is assumed to be

$$Eu_{it} = \theta_i E\left[q \mid I_i(t)\right] - p,$$

where θ_i is the willingness to pay for a movie's quality and it is a realization from $\theta_i \sim U\left[\underline{\theta}, \overline{\theta}\right]$, $E\left[q \mid I_i(t)\right]$ is the expected entertainment value of a new movie perceived by consumer i at time t, based on the individual information set, $I_i(t)$ and p is the price of watching a movie in the theatre. The whole market size is normalized to 1 and in period 1, only a fraction (γ) of consumers have chances to enter the market, but all consumers can enter the market in period 2, as long as they are informed about the movie's arrival. Here, γ being small has the same intuition as **Assumption 1** which means that long-run market is important enough for studios. So the market share for the studio in period 1 will be determined by three factors: how many consumers enter the market (γ), how many consumers are covered by the ads $(\varphi_1(a_1))$ and how many consumers are convinced by the ads $(\tau_1(a_1))$. Then, we have

$$\delta_1 = \gamma\varphi_1(a_1)\,\tau_1(a_1) = \delta_1(a_1)$$

For simplicity, I assume the studio only advertise in period 1, then we have

$$\pi\left(a_1, q\right)=\varphi\left(a_1\right)\left[\gamma \tau_1\left(a_1\right)+\left(1-b \gamma \tau_1\left(a_1\right)\right) \tau_2\left(\delta_1, q\right)\right]$$

and

$$\begin{aligned} \frac{\partial^2 \pi\left(a_1, q\right)}{\partial a_1 \partial q} &= \left[\varphi^{'}\left(a_1\right)-\gamma\left(\varphi^{'}\left(a_1\right) \tau_1\left(a_1\right)+\varphi\left(a_1\right) \tau_1^{'}\left(a_1\right)\right)\right] \frac{\partial \tau_2}{\partial q} \\ &+\varphi\left(a_1\right)\left(1-\gamma \tau_1\left(a_1\right)\right) \frac{\partial^2 \tau_2}{\partial a_1 \partial q} \end{aligned} \tag{14}$$

Here, I will discuss several cases about how WOM in period2 is impacted by market share in period 1 and movie's true quality ($\tau_2\left(\delta_1, q\right)$ as a function of δ_1 and q).

Case 1: in period 2, q is only learned through WOM and σ_w^2 is independent of δ_1.

In this case, $\tau_2=\tau_2(q)$, so $\frac{\partial^2 \tau_2\left(\delta_1, q\right)}{\partial \delta_1 \partial q}=0$. When consumers don't use any information from ads intensity to update their beliefs in period 1, then $\tau_1^{'}\left(a_1\right)=0$, and $\frac{\partial^2 \pi\left(a_1, q\right)}{\partial a_1 \partial q}=\varphi^{'}\left(a_1\right)\left[1-\gamma \tau_1\left(a_1\right)\right] \frac{\partial \tau_2}{\partial q} \geq 0$. When consumers do use information from ads intensity to update their beliefs in period 1, $\tau_1=\tau_1\left(a_1\right)$ should be a nondecreasing function of a_1 and $\frac{\partial^2 \pi\left(a_1, q\right)}{\partial a_1 \partial q}=\left[\varphi^{'}\left(a_1\right)\left(1-\gamma \tau_1\left(a_1\right)\right)-\gamma \varphi\left(a_1\right) \tau_1^{'}\left(a_1\right)\right] \frac{\partial \tau_2}{\partial q} \geq 0$ if γ is small enough.

Case 2: q is learned both through WOM and δ_1; σ_w^2 is independent of δ_1

In this case, $\tau_2=\tau_2\left(\delta_1, q\right)$, and δ_1 only reflects the studio's advertising spending a_1 and therefore q_s received by the studio. σ_w^2 being independent of δ_1 means that the accuracy of the WOM is the same when even more people watch the movie in period 1. With the assumption that $g_0(q)$, $f\left(q_s \mid q\right)$ and $f\left(q_{iw} \mid q, \sigma_w^2\right)$ are normal distributions, the expected quality perceived by consumer i in period 2 is a linear combination of $\overline{q}_{iw}$, $q_s = \Delta_1^{-1}(\delta_1)$ and $\overline{q}$:

$$E_{i2}^c(q) = E_{i1}^c[q \mid \delta_1] + \beta_q^c(1)\left(\overline{q}_{iw} - E_{i1}^c[q \mid \delta_1]\right)$$

where $E_{i1}^c[q \mid \delta_1] = E_1^s(q) = \overline{q} + \beta_q(0)\left(\Delta_1^{-1}(\delta_1) - \overline{q}\right)$

Then consumer i chooses to watch a new movie when

$$\begin{aligned} \theta_i E q_2 - p &\geq 0 \\ \theta_i &\geq \frac{p}{Eq_2} = h(q_{iw}, \delta_1). \end{aligned}$$

In this case, $\tau_2(\delta_1, q)$ can be written as

$$\begin{aligned} \tau_2(\delta_1, q) &= \int_{q_{iw}} \frac{\overline{\theta} - h(q_{iw}, \delta_1)}{\overline{\theta} - \underline{\theta}} \phi\left(q_{iw} \mid q, \sigma_w^2\right) d_{q_{iw}} \\ \tau_2^q &= \frac{\partial \tau_2(\delta_1, q)}{\partial q} = \int_{q_{iw}} \frac{\overline{\theta} - h(q_{iw}, \delta_1)}{\overline{\theta} - \underline{\theta}} \phi\left(q_{iw} \mid q, \sigma_w^2\right) \left(\frac{q_{iw} - q}{\sigma_w^2}\right) d_{q_{iw}} \\ \tau_2^{q\delta} &= \frac{\partial^2 \tau_2(\delta_1, q)}{\partial \delta_1 \partial q} = \int_{q_{iw}} \frac{-h^\delta(q_{iw}, \delta_1)}{\overline{\theta} - \theta} \phi\left(q_{iw} \mid q, \sigma_w^2\right) \left(\frac{q_{iw} - q}{\sigma_w^2}\right) d_{q_{iw}} \end{aligned}$$

Here, $h^\delta(q_{iw}, \delta_1) = \frac{\partial h(q_{iw}, \delta_1)}{\partial \delta_1} < 0$. Since

$$h^{q_w \delta}(q_{iw}, \delta_1) = \frac{\partial^2 h(q_{iw}, \delta_1)}{\partial \delta_1 \partial q_{iw}} = \frac{p \beta_q^c(1)\left(1 - \beta_q^c(1)\right) E q_1^{'}(\delta_1)}{E q_2^3} > 0,$$

so $\tau_2^{q\delta} > 0$. When consumers don't use any information from ads intensity to update their beliefs in period 1, then $\tau_1^{'}(a_1) = 0$ and

$$\begin{aligned} \frac{\partial^2 \pi(a_1, q)}{\partial a_1 \partial q} &= \varphi^{'}(a_1)\left[1 - \gamma \tau_1(a_1)\right] \frac{\partial \tau_2}{\partial q} + \varphi(a_1)\left(1 - \gamma \tau_1(a_1)\right) \frac{\partial^2 \tau_2}{\partial a_1 \partial q} \\ &= \left[\varphi^{'}(a_1) \frac{\partial \tau_2}{\partial q} + \varphi(a_1) \frac{\partial^2 \tau_2}{\partial a_1 \partial q}\right] \left[1 - \gamma \tau_1(a_1)\right] > 0 \end{aligned} \quad (15)$$

When consumers do use information from ads intensity to update their beliefs in period 1, $\tau_1 = \tau_1(a_1)$ and $\frac{\partial^2 \pi(A_1,q)}{\partial A_1 \partial q} \geq 0$, as long as γ is small enough.

Case 3: q is learned only through WOM. σ_w^2 is decreasing in δ_1

In this case, $\beta_q^c(1)$ being increasing in δ_1 means that more weight is put on the information from WOM when more people watch the movie in period 1. Therefore δ_1 complement the role of q more for good movies. Then A2 should be easier to be satisfied in this case than in case 2

Case 4: q is learned only through WOM and δ_1. σ_w^2 is decreasing in δ_1

The combination of Case 2 and Case 3.

A.2: Computation of Jacobian Matrix

The Jacobian matrix in equation 9 is

$$
\begin{aligned}
J_{(j,Z\to Y)} &= \frac{\partial \zeta^{-1}\left(Y_j|X_j,\Theta\right)}{\partial Y_j} \\
&= \begin{bmatrix}
\frac{\partial q_{js}}{\partial a_{j1}} & \frac{\partial q_{js}}{\partial \delta_{j1}} & \frac{\partial q_{js}}{\partial a_{j2}} & \frac{\partial q_{js}}{\partial \delta_{j2}} \\
\frac{\partial \eta_{j1}}{\partial a_{j1}} & \frac{\partial \eta_{j1}}{\partial \delta_{j1}} & \frac{\partial \eta_{j1}}{\partial a_{j2}} & \frac{\partial \eta_{j1}}{\partial \delta_{j2}} \\
\frac{\partial q_{j}}{\partial a_{j1}} & \frac{\partial q_{j}}{\partial \delta_{j1}} & \frac{\partial q_{j}}{\partial a_{j2}} & \frac{\partial q_{j}}{\partial \delta_{j2}} \\
\frac{\partial \eta_{j2}}{\partial a_{j1}} & \frac{\partial \eta_{j2}}{\partial \delta_{j1}} & \frac{\partial \eta_{j2}}{\partial a_{j2}} & \frac{\partial \eta_{j2}}{\partial \delta_{j2}}
\end{bmatrix} \\
&= \begin{bmatrix}
\frac{\partial q_{js}}{\partial a_{j1}} & 0 & 0 & 0 \\
\frac{\partial \eta_{j1}}{\partial a_{j1}} & \frac{\partial \eta_{j1}}{\partial \delta_{j1}} & 0 & 0 \\
\frac{\partial q_{j}}{\partial a_{j1}} & \frac{\partial q_{j}}{\partial \delta_{j1}} & \frac{\partial q_{j}}{\partial a_{j2}} & 0 \\
\frac{\partial \eta_{j2}}{\partial a_{j1}} & \frac{\partial \eta_{j2}}{\partial \delta_{j1}} & \frac{\partial \eta_{j2}}{\partial a_{j2}} & \frac{\partial \eta_{j2}}{\partial \delta_{j2}}
\end{bmatrix}
\end{aligned}
$$

To get the Jacobian term, I only need to get the following terms:

$$\frac{\partial q_{js}}{\partial a_{j1}} = \frac{1}{\frac{\partial A_1\left(x_j,q_{js}\right)}{\partial q_{js}}} = \frac{\partial H(a_{j1},x_j)}{\partial a_{j1}},$$

$$\frac{\partial \eta_{j1}}{\partial \delta_{j1}} = \frac{1}{\frac{\partial \Delta_1\left(x_j,a_{j1},\eta_{j1}\right)}{\partial \eta_{j1}}} = \frac{1}{\varphi_1(a_{j1})\tau_1\left(a_{j1},\eta_{j1}\right)\left(1-\tau_1\left(a_{j1},\eta_{j1}\right)\right)} = \frac{1}{\delta_{j1}\left(1-\tau_1\left(a_{j1},\eta_{j1}\right)\right)},$$

$$\frac{\partial q_{j}}{\partial a_{j2}} = \frac{1}{\frac{\partial A_2\left(x_j,a_{j1},\delta_{j1},q_j\right)}{\partial q_j}} = \frac{1}{\frac{\partial E_2(\tau_2)/\partial q_j}{\lambda E_2(\tau_2)}} = \frac{\exp(\lambda(\mu * a_{j1}+a_{j2}))}{Mp\partial E_2(\tau_2)/\partial q_j},$$

$$\frac{\partial \eta_{j2}}{\partial \delta_{j2}} = \frac{1}{\frac{\partial \Delta_2\left(x_j,a_{j1},\delta_{j1},a_{j2},q_j,\eta_{j2}\right)}{\partial q_j}} = \frac{1}{\Psi(a_{j1},a_{j2},\delta_{j1})\int \tau_2(\eta_2|\cdot)(1-\tau_2(\eta_2|\cdot))d_{f(E^c[q|I(2)]|q)}}.$$

A.3: Tables

Table 1: Summary Statistics of Main Variables

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

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: Estimation Results for Advertising Spending Regressions

Dependent Variable	a1		a2	
	Non-sequel Movies	Sequel Movies	Non-sequel Movies	Sequel Movies
bot1			1.17***	0.62***
			(0.063)	(0.108)
budget	0.12***	0.08***	0.01*	0.02
	(0.007)	(0.012)	(0.004)	(0.011)
critic	0.08***	0.04	0.06***	0.08***
	(0.011)	(0.028)	(0.007)	(0.022)
summer	-0.32	0.68	0.01	-0.49
	(0.401)	(0.845)	(0.230)	(0.639)
holiday	1.60***	3.40**	2.65***	3.79***
	(0.611)	(1.409)	(0.350)	(1.077)
# of competitors	-0.47***	0.58*	0.02	0.45*
	(0.156)	(0.34)	(0.093)	(0.263)
runtime	0.03**	0.03	0.02***	0.08***
	(0.015)	(0.028)	(0.008)	(0.021)
major	2.27***	2.93**	-0.25	-1.03
	(0.578)	(1.171)	(0.332)	(0.885)
Mini-major	1.3**	0.60	-0.23	-1.12
	(0.713)	(1.665)	(0.409)	(1.263)
Action	0.81	1.78*	-0.70**	0.09
	(0.532)	(1.077)	(0.306)	(0.815)
comedy	1.81***	1.74**	0.54**	0.30
	(0.416)	(0.915)	(0.239)	(0.689)
drama	0.06	-1.32	0.15	-1.03
	(0.539)	(4.044)	(0.309)	(3.107)
family	0.06	4.21**	0.10	3.01*
	(0.915)	(2.088)	(0.525)	(1.571)
horror	-1.26**	0.30	-0.68*	1.20
	(0.624)	(1.282)	(0.362)	(0.964)
mpaa_G	0.79	-3.05	1.79**	-0.04
	(1.320)	(2.584)	(0.764)	(1.944)
mpaa_PG	2.22**	-0.91	0.78*	0.66
	(0.790)	(1.885)	(0.458)	(1.421)
mpaa_PG13	1.17**	3.19***	-1.11	0.67
	(0.409)	(1.066)	(0.238)	(0.805)
Constant	-0.79	0.15	-3.60***	-10.53***
	(1.635)	(3.437)	(0.945)	(2.587)
Observations	632	111	632	111
R-squared	0.59	0.71	0.65	0.82
Adj. R-squared	0.58	0.65	0.64	0.79

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3: Estimation Results for Box office Performance Regressions

Dependent Variable	log(bot1)		log(bot2)	
	Non-sequel Movies	Sequel Movies	Non-sequel Movies	Sequel Movies
log(a1)	0.41***	1.13**		
	(0.111)	(0.465)		
log(a1+a2)			0.90***	1.38***
			(0.102)	(0.391)
f(qs)	0.03***	-0.06*	-0.03***	-0.04**
	(0.010)	(0.034)	(0.008)	(0.022)
f(q)			0.07***	0.02
			(0.009)	(0.018)
bot1			0.31***	0.11***
			(0.014)	(0.018)
budget	0.01***	0.00	-0.01***	-0.00
	(0.001)	(0.003)	(0.001)	(0.002)
critic	0.01***	0.01***	0.00***	0.01***
	(0.002)	(0.004)	(0.002)	(0.004)
summer	-0.03	-0.01	0.02	0.27***
	(0.055)	(0.110)	(0.046)	(0.089)
holiday	0.05	-0.36*	0.22***	0.08
	(0.084)	(0.199)	(0.072)	(0.168)
# of competitors	-0.14***	-0.14***	0.04*	-0.04
	(0.022)	(0.049)	(0.019)	(0.041)
runtime	0.00	0.00	0.00***	0.00
	(0.002)	(0.004)	(0.002)	(0.003)
major	0.16*	-0.17	-0.05	0.21
	(0.086)	(0.195)	(0.071)	(0.140)
mini major	0.18*	-0.17	0.01	0.42**
	(0.100)	(0.222)	(0.084)	(0.175)
action	0.11	-0.24	-0.05	-0.06
	(0.073)	(0.151)	(0.061)	(0.119)
comedy	0.03	-0.16	0.01	0.19
	(0.059)	(0.149)	(0.049)	(0.113)
drama	-0.07	-1.66***	-0.11*	-0.82*
	(0.073)	(0.523)	(0.062)	(0.423)
family	0.00	-0.07	-0.00	0.40
	(0.125)	(0.313)	(0.105)	(0.249)
horror	0.39***	0.12	0.06	0.17
	(0.086)	(0.168)	(0.074)	(0.137)
mpaa_G	0.15	0.05	0.43***	-0.18
	(0.175)	(0.359)	(0.142)	(0.274)
mpaa_PG	1.12	-0.22	0.08	-0.36*
	(0.106)	(0.245)	(0.086)	(0.194)
mpaa_PG13	0.13**	-0.19	0.02	-0.14
	(0.055)	(0.172)	(0.044)	(0.125)
Constant	-2.03***	-2.06**	-3.41***	-2.56***
	(0.281)	(0.847)	(0.224)	(0.634)
Observations	632	111	632	111
R-squared	0.53	0.69	0.80	0.88
Adj. R-squared	0.51	0.62	0.80	0.85

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4: Short-run Vs Long-run Market

	# of obs	Mean	Std. Dev.	Min	Max
All movies in the sample:					
Market share for opening week	632	0.82%	0.70%	0.04%	4.18%
Market share for post-release weeks	632	1.95%	2.08%	0.02%	15.39%
Pre-release ads/Total ads>95%:					
Total ads spending	20	11.01	4.06	2.77	19.31
Total BOR	20	6.85	3.77	1.07	14.38
Total BOR/ Total ads spending	20	63.29%	30.77%	17.50%	126.46%
Pre-release ads/Total ads	20	97.02%	1.54%	95.02%	99.84%
Opening week BOR/Total BOR	20	48.30%	7.94%	32.59%	63.22%
Budget	20	30.88	18.13	2.65	90.00
Metacritic (1-100)	20	32	12	17	57
IMDB user rating (1-10)	20	5	1	2	7
Pre-release ads/Total ads <60%:					
Total ads spending	34	30.51	13.42	7.34	55.90
Total BOR	34	122.13	84.00	13.54	339.72
Total BOR/ Total ads spending	34	389.67%	193.08%	49.70%	964.69%
Pre-release ads/Total ads	34	53.17%	5.44%	35.30%	59.98%
Opening week BOR/Total BOR	34	23.08%	6.97%	12.20%	43.04%
Budget	34	57.21	42.58	3.00	165.00
Metacritic (1-100)	34	64	15	31	90
IMDB user rating (1-10)	34	7	1	5	8

Table 5: Approximated Inversed Advertising Policy Function

a1 \ X	0	0.20	0.40	0.60	0.80	1.00	1.20	1.40	1.60	1.80
0.68	-2.67	-3.61	-4.37	-4.99	-5.50	-5.93	-6.32	-6.70	-7.10	-7.55
4.96	1.03	-0.35	-1.60	-2.73	-3.77	-4.73	-5.66	-6.55	-7.45	-8.37
9.24	4.19	2.52	0.96	-0.50	-1.89	-3.21	-4.49	-5.74	-6.97	-8.21
13.52	6.90	5.06	3.32	1.65	0.04	-1.51	-3.02	-4.49	-5.95	-7.40
17.80	9.24	7.32	5.46	3.67	1.94	0.25	-1.40	-3.02	-4.62	-6.21
22.08	11.30	9.32	7.41	5.55	3.74	1.97	0.23	-1.48	-3.17	-4.85
26.35	13.12	11.12	9.17	7.27	5.41	3.59	1.79	0.02	-1.73	-3.47
30.63	14.76	12.74	10.77	8.83	6.93	5.07	3.24	1.42	-0.38	-2.17
34.91	16.24	14.21	12.21	10.25	8.32	6.42	4.54	2.68	0.84	-0.98
39.19	17.60	15.54	13.52	11.53	9.57	7.64	5.73	3.84	1.96	0.10

Table 6.1: Estimated Parameters for Utility Function

	Estimates	Standard Error
Coefficients for Observed Characteristics (γ)		
Budget	0.064***	0.0002
Critic	0.058***	0.0006
Summer	-0.292	0.3211
Holiday	0.826	1.0756
# of competitors	-0.030***	0.0016
Runtime	0.221***	0.0508
Major	0.129***	0.0123
Mini-major	0.071***	0.0134
Action	0.069***	0.0068
Comedy	1.075*	0.6172
Drama	-0.125	1.0835
Family	-0.038	4.8666
Horror	-0.052***	0.0099
MPAA_G	0.076	0.1289
MPAA_PG	0.149***	0.0395
MPAA_PG13	0.092***	0.0037
Time dummy (α)	0.784***	0.0093
Demand shock variance in period 1 ($\sigma^2_{\eta_1}$)	0.351***	0.0022
Demand shock variance in period 2 ($\sigma^2_{\eta_2}$)	0.337***	0.0029

Table 6.2: Estimated Parameters for Information Learning

	Estimates	Standard Error
Mean of quality ($\bar{q}$)	-0.317***	0.0645
Variance of quality (σ^2_q)	0.339***	0.0084
Variance of noisy signal of quality (σ^2_s)	3.501**	1.5168
Word of Mouth variance (adjusted) ($\frac{\sigma^2_w}{\rho}$)	0.023**	0.0106

Table 6.3: Estimated Parameters for Advertising Reach Function

	Estimates	Standard Error
Reach efficiency parameter (λ)	0.131***	0.0269
Advertising depreciation parameter (μ)	0.316***	0.0005

Note: *** p<0.01, ** p<0.05, * p<0.1

Table 7-1: Data Vs Model

Cells	Data	Pre-release Advertising	Post-release Advertising	Opening Week Performance	Post-release Weeks Performance
Cell 1	126	132.61	124.69	139.41	141.83
Cell 2	127	147.65	110.49	135.26	134.05
Cell 3	126	122.61	127.03	122.98	134.51
Cell 4	127	109.83	141.55	128.94	112.29
Cell 5	126	119.29	128.24	105.41	109.32
X^2 Stat.		6.37	4.02	5.92	7.15

Note: cells for different response variables have different ranges;

$X^2_{4,0.10} = 7.78$

$X^2_{4,0.05} = 9.49$

Table 7-2: Data Vs Model: Pre-release Advertising Conditional on Observed Attributes

		X									
		<0.70		[0.70,0.85]		[0.85,0.97]		[0.97,1.15]		>1.15	
		Data	Model	Data	Model	Data	Model	Data	Model	Data	Model
	<9.77	78	70.00	29	35.68	14	18.34	4	6.94	1	1.65
	[9.77,13.77]	39	39.88	39	46.16	33	35.58	15	22.08	1	3.96
a1	[13.77,17.18]	8	12.83	38	29.60	42	36.56	32	32.32	6	11.30
	[17.18,21.44]	1	3.13	17	13.23	29	27.41	44	39.78	36	26.30
	>21.44	0	0.16	4	2.33	8	8.11	32	25.88	82	82.80
	X2 stat		4.36		7.01		2.12		5.42		8.53

$X^2_{4,0.10} = 7.78$

$X^2_{4,0.05} = 9.49$

Table 8: Signaling VS Reaching Role of Advertising

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

Note: All advertising spending is in billions, and all numbers are calculated based on the target sample.

A.4: Figures

Figure 1: Pre-release and Post-release Advertising Spending VS Production Budget

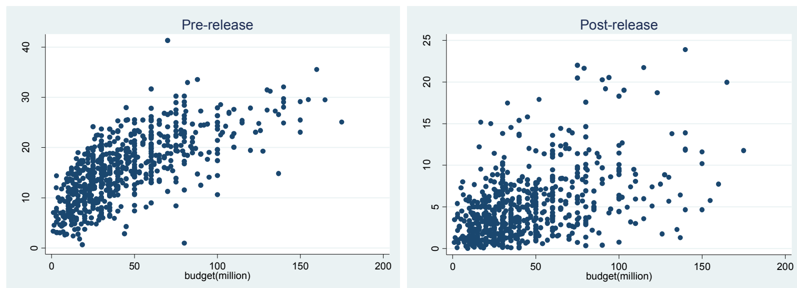
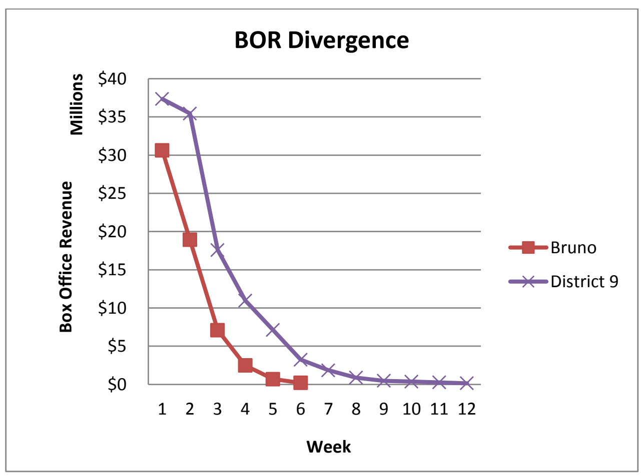


Figure 2: Sales trends diverge over time after release



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

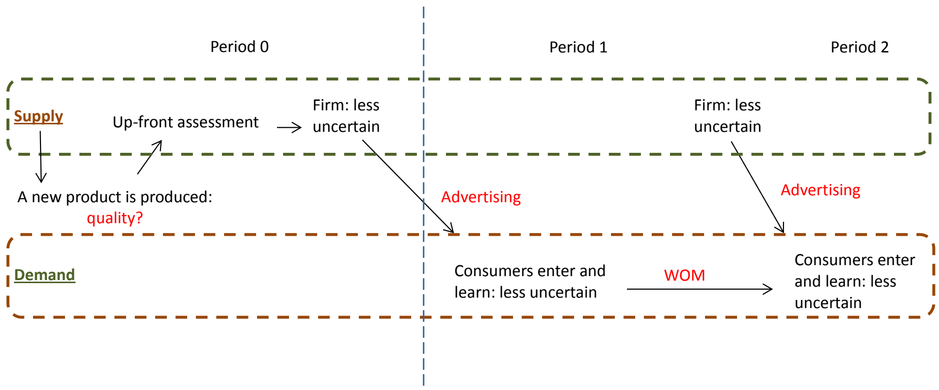


Figure 4: The Importance of Long-run Market

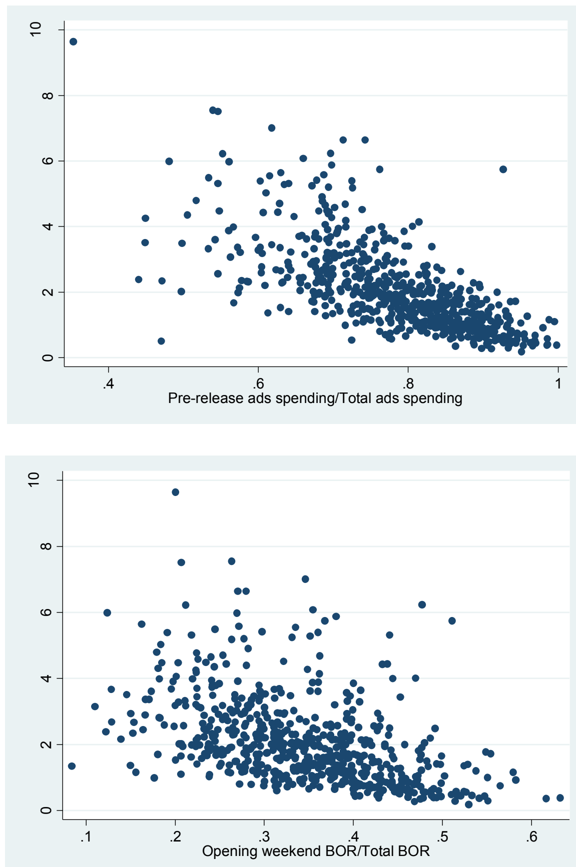
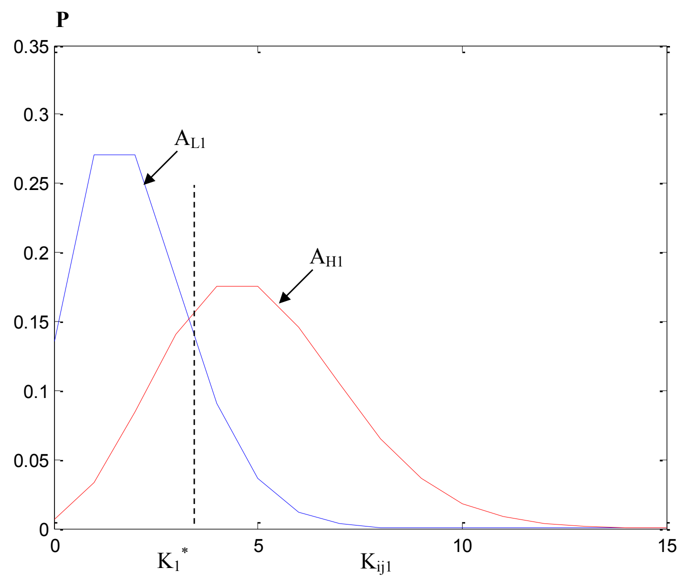


Figure 5: Distribution of K_{ijt}



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 6-1: advertising spending function a=f(qs) when roh=0.05

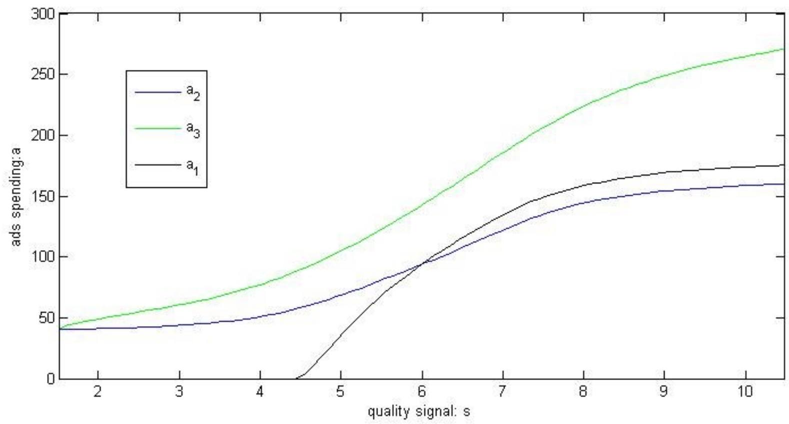


Figure 6-2: advertising spending function a=f(qs) when roh=0.1

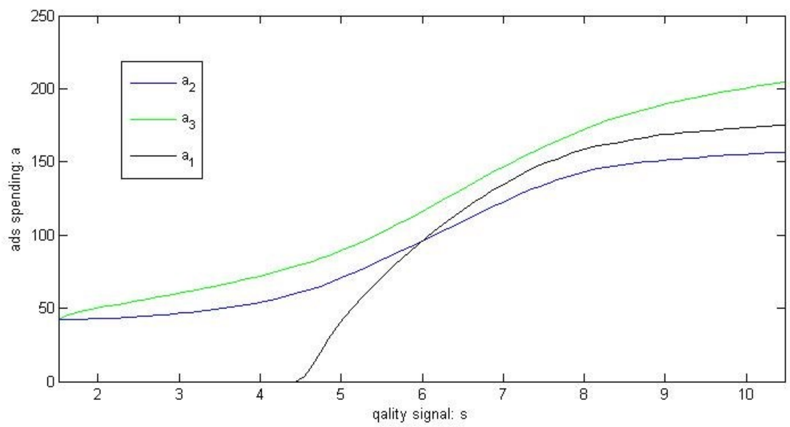


Figure 7: Inverse Pre-release Advertising Policy Function

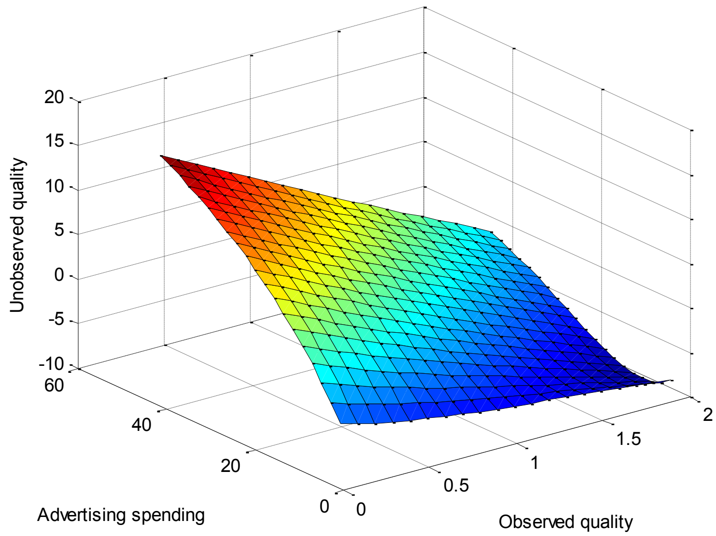


Figure 8: Advertising Strategy under Different Information Structure

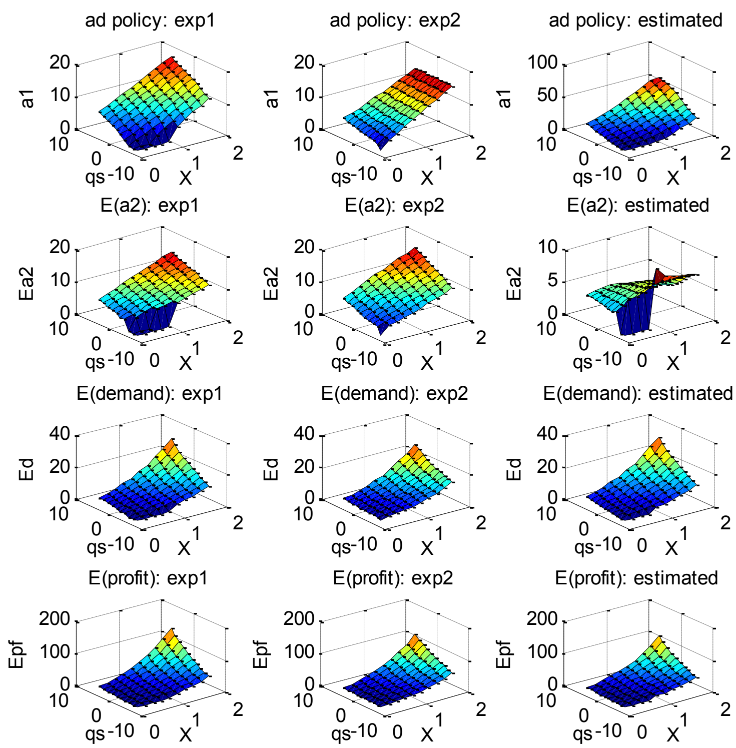
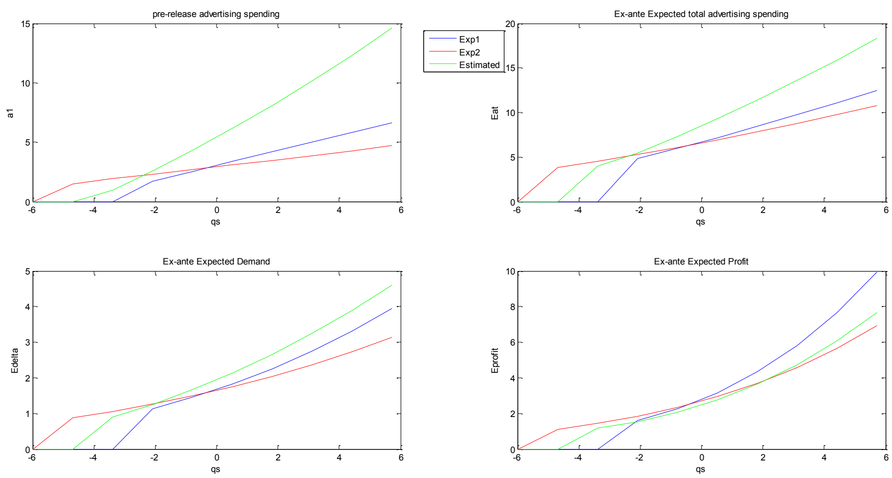


Figure 9: Comparison of Three Cases (Conditional on X)



References

[1] Ackerberg, Daniel A, "Empirically Distinguishing Informative and Prestige Effects of Advertising," *RAND Journal of Economics*, 2001, 32(2), 316-33;

[2] Ackerberg, Daniel A, "Advertising, Learning and consumer choice in experience good market: an empirical examination," *International Economic Review*, 2003, 44(3), 1007-1040;

[3] Aderson, Simon and Renault, "Advertising Content," *American Economic Review*, 2006, 96(1), 93-113;

[4] Andrew, Donald, "Chi-square Diagnostic Tests for Econometric Models: Theory," *Econometrica*, 1988, 56(6), 1419-1453;

[5] Athey, Susan, "Monotone Comparative Statics under Uncertainty," The Quarterly Journal of Economics, 2002, 117(1), 187-223;

[6] Bagwell, Kyle, "The Economics of Advertising, Introduction," The Economics of Advertising, 2001, Edward Elgar Publishing;

[7] Basuroy, Suman, Kalpesh Kaushik Desai, and Debabrata Talukdar, "An Empirical Investigation of Signaling in the Motion Picture Industry," *Journal of Marketing Research*, 2006, 43(2), 287-295;

[8] Butters, Gerard, "Equilibrium Distributions of Sales and Advertising Prices," *Review of Economic Studies*, 1977, 44, 465–91;

[9] Byzalov, Dmitri and Ron Shachar, "The Risk Reduction Role of Advertising," *Quantitative Marketing and Economics*, 2004, 2(4), 283-320;

[10] Ching, Andrew T., "A Dynamic Oligopoly Structural Model For The Prescription Drug Market After Patent Expiration," *International Economic Review*, 2010, 51(4),1175-1207;

[11] Ching, Andrew T., "Consumer Learning and Heterogeneity: Dynamics of Demand for Prescription Drugs After Patent Expiration," *International Journal of Industrial Organization*, 2010, 28(6), 619-638;

[12] Conlon, Christopher T., "A Dynamic Model of Costs and Margins in the LCD TV Industry," 2010;

[13] De Vany, Arthur and Cassey Lee., "Quality Signals in Information Cascades and the Dynamics of the Distribution of Motion Picture Box Office Revenues," *Journal of Economic Dynamic and Control*, 2001, 25, 593-614;

[14] Eckstein, Z., D. Horsky, Y. Raban, "An empirical dynamic model of brand choice," Working paper, 1988, University of Rochester;

[15] Einav, Liran, "Seasonality and Competition in Time: An Empirical Analysis of Release Date Decisions in the U.S. Motion Picture Industry," *RAND Journal of Economics*, 2007, 38(1), 127-145;

[16] Elberse, Anita, Bharat Anand, "The Effectiveness of Pre-release Advertising for Motion Pictures: An Empirical Investigation Using a Simulated Market," *Information Economics and Policy*, 2007, 19, 319-343;

[17] Erdem, Tulin and Michael P. Keane, "Decision-making Under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets," *Marketing Science*, 1996, 15(1), 1-20;

[18] Erdem, Tulin, Michael P. Keane and Baohong Sun, "A Dynamic Model of Brand Choice When Price and Advertising Signal Product Quality," *Marketing Science*, 2008, 27(6), 1111–1125;

[19] Grossman, G., and C. Shapiro, "Informative Advertising with Differentiated Products," *The Review of Economic Studies*, 1984, 51, 63–81;

[20] Hertzendorf, Mark N., "I'm Not a High-Quality Firm-But I Play One on TV," *The RAND Journal of Economics*, 1993, 24(2), 236-247;

[21] Horstman, I.J. and MacDonald, G., "Advertising as a Signal of Product Quality: Compact Disc Players," *International Journal of Industrial Organization*, 2003, 21, 317-345;

[22] Joo, Hailey Hayeon, "Social Learning and Optimal Advertising in the Motion Picture Industry," University of Pennsylvania, Dissertation,January 1, 2008;

[23] Joshi, Amit M. and Dominique M. Hanssens, "Movie Advertising and the Stock Market Valuation of Studios: A Case of "Great Expectations"?," *Marketing Science*, 2009, 28(2), 239-250;

[24] Kennedy, Peter W., "Word-of Mouth Communication and Prices as a Signal of Quality," *The Economic Record*, 1994, 70, 373-380;

[25] Kihlstrom, Richard E. and Michael H. Riordan, "Advertising as a Signal," *The Journal of Political Economy*, 1984, 92(3), 427-450;

[26] Linnemer, Laurent, "Price and Advertising as Signals of Quality when Some Consumers are Informed," *International Journal of Industrial Organization*, 2002, 20, 931-947;

[27] Linnemer, Laurent, "Dissipative Advertising Signals Quality even without Repeat Purchases," 2008, CESifo Working Paper Series No. 2310;

[28] Marich, Robert, "Marketing to Moviegoers: A Handbook of Strategies and Tactics", 2009, second edition;

[29] Matthews, Steven A. and Leonard J. Mirman, "Equilibrium Limit Pricing: The Effects of Private Information and Stochastic Demand," 1983, *Econometrica*, 51(4), 981-996;

[30] Milgrom, P., and J. Roberts, "Price and Advertising Signals of Product Quality," *Journal of Political Economy*, 1986, 94, 796–21;

[31] Moretti, Enrico, "Social Learning and Peer Effects in Consumption: Evidence from Movie Sales," *Review of Economic Studies,* 2011, 78, 356–393;

[32] Moul, Charles C., "Measuring Word of Mouth's Impact on Theatrical Movie Admissions," *Journal of Economics and Management Strategy*, 2007, 6(4), 859-892;

[33] Narayanan, Sridhar, Puneet Manchanda and Pradeep K., "Temporal Differences in the Role of Marketing Communication in New Product Categories," *Journal of Marketing Research*, 2005, 42, 278-290;

[34] Nelson, P., "Information and Consumer Behavior," *Journal of Political Economy*, 1970 78, 311-29;

[35] Nelson, P., "Advertising as Information," *Journal of Political Economy*, 1974, 82, 729–753;

[36] Rennhoff, Adam D. and Kenneth C. Wilbur, "The Effectiveness of Post-Release Movie Advertising," 2008, Marshall School of Business Working Paper No. MKT 08-08;

[37] Roberts, John H. and Glen L. Urban, "Modeling Multiattribute Utility, Risk, and Belief Dynamics for New Consumer Durable Brand Choice," *Management Science*, 1988, 34(2), 167-185;

[38] Santugini, Marc, "Observational Learning in the Motion Picture Market," Department of Economics, University of Virginia, January 11, 2007;

[39] Sovinsky Goeree M., "Limited information and advertising in the U.S. personal computer industry," *Econometrica*, 2008, 76(5), 1017–1074;

[40] Su, Chelin and Kenneth L. Judd, "Constrained Optimization Approaches to Estimation of Structural Models," *Econometrica*, 2012, 80(5), 2213–2230;

[41] Su, Chelin, "Estimating Discrete-Choice Games of Incomplete Information: Simple Static Examples," *Quantitative Marketing and Economics,* April 2014, forthcoming.

[42] Thomas, L., Shane, S. and K. Weiglet, "An Empirical Examination of Advertising as a Signal of Quality," *Journal of Economic Behavior* and *Organization*, 1998, 37, 415-30;

[43] Vitorino, Maria Ana, "Empirical Entry Games with Complementarities: an Application to The Shopping Center Industry," 2011;