Strategic Advertising and Release in the Movie Industry

Weichen Yan[‡]

Abstract

Setting the release date and advertising for a blockbuster movie are the most important decisions a studio makes in movie distribution. I propose a model which combines studios' release date (entry) and advertising choices under incomplete information. Studios optimize by setting release dates followed by advertising budget, and have payoffs structurally determined and dependent on demand. I adopt an estimation method which accounts for advertising and choice set endogeneities, the latter leading to a selection bias if not corrected. Model estimates, in particular the effect of advertising, are significantly different from the traditional estimates. I show that ignoring advertising and/or choice set endogeneity significantly underestimates the impact of advertising on box office. In light of the recent Disney-Fox merger, I simulate how this event will impact the movie industry.

Keywords: motion pictures, endogenous choice set, demand estimation, entry game, advertising, semiparametric estimation

[*]I am indebted to Luis Cabral, Isabelle Perrigne and Quang Vuong for their guidance and support. I would like to thank Xinyu Cao, Giovanni Compiani, Timothy Christensen, Liran Einav, Tulin Erdem, Anindya Ghose, Yufeng Huang, Konrad Menzel, Steven Shugan, K. Sudhir and Raluca Ursu for helpful comments and suggestions. I am also grateful to seminar participants at NYU Econometrics Lunch, NYU Stern IO Workshop, NYU Stern Marketing Seminar and conference participants at INFORMS Marketing Science Conference for their valuable comments and insights. I would like to thank NYU-CRATE for its financial support. All errors are my own.

[‡]Department of Economics, New York University. Email: wy462@nyu.edu

1 Introduction

Distribution is one of the most important stages of the movie industry's value chain. Two strategic decisions are key in motion picture distribution: advertising and release date, as pointed out by Eliashberg, Elberse and Leenders (2006). In particular, when it comes to "blockbuster" movies, that is, high-budget, star-studded movies with large advertising campaigns, distributors must choose a release date with incomplete information regarding rival blockbusters' unobservable characteristics and release dates. Moreover, understanding the impact of advertising on box-office performance and therefore selecting the optimal advertising level is a challenging question due to a number of endogeneity issues in estimation.

As an illustration, two big-budget action movies stood out in summer 2015 : Paramount's *Terminator: Genesys* (released July 1) and Disney's *Ant Man* (released July 17). A natural question is whether *Ant Man* would have performed better if Disney had released it during the 4th of July weekend. On the one hand, potential demand would likely be higher, but on the other hand *Ant Man* would have to go head-to-head against *Terminator*. Later in summer of 2015, Fox released action blockbuster *Fantastic Four*, while Sony released action blockbuster *Pixels* (released July 24). *Fantastic Four* (released Aug 7) grossed \$149 million at the box office with an advertising budget of \$25.7 million. By contrast, *Pixels* grossed \$80 million at the box office with an advertising budget of \$33 million. The combination of higher advertising and lower box office may wrongly suggest that advertising has no effect, or potentially a negative effect on box office. However, one must consider the possibility that Sony's large advertising budget could reflect the distributor's private knowledge regarding the appeal or "quality" of *Pixels*. Sony might compensate a lower quality movie by a larger amount of advertising to maximize attendance.

The above examples suggest a variety of interesting research questions: (i) how a studio's private knowledge of its movie quality affects its release date, (ii) how much a movie's performance is due to the movie being a sequel to a well-known series, (iii) how being released during a high-demand weekend affects movie performance, and (iv) how distributors' private information regarding their movie's appeal affects their choice of advertising budget. These are some of the questions I address in this paper. Specifically, I develop a structural model of supply and demand. On the supply side, I model distributors first making release-date decisions in an incomplete information setup, and then, given all the announced release dates, deciding on their blockbuster's advertising budget. Each distributor's decision is a function of private information regarding its movie quality.[1] On the demand side, I consider a discrete choice model similar to Einav (2007), my demand model however incorporates the endogenous nature of advertising budgets and release dates.

Failing to account for the endogeneity of release date and advertising decisions may lead to biased and inconsistent estimates of key parameters, as shown by Perrigne, Vuong and Yan (2019). I address these issues by explicitly modeling the entry (release date) and advertising decisions, and by employing a combination of instrumental variables and control function methods. I show that the estimates of the effect on box office of advertising and other important variables characterizing the movie are significantly different once these endogeneities are accounted for. For instance, I estimate that everything else constant, the positive impact of being a sequel on box office is smaller when taking into account advertising and/or choice set endogeneity, compared to assuming exogenous advertising and/or release. Intuitively, a sequel to a well-known series is more likely to be released during a high-demand period. As such, a naive approach that takes supply as given leads to a positive bias in the estimation of the sequel dummy.

Perhaps more surprisingly, I also find that distributors tend to compensate for low movie quality by setting high advertising budgets. Indeed, estimating the model with endogenous advertising and without, I find a negative bias in its coefficient. Consider two movies, one with high quality, one with low quality. The low-quality movie spends more on advertising. It does better at the box office, but not much better than the

[1] I allow for the possibility that a distributor has no blockbuster available, in which case the release-date and advertising decisions are trivial.

high-quality movie, which spent less in advertising. A naive approach that does not take endogeneity into account assigns a low effect to advertising, ignoring the fact that the high-advertising movie could be a low-quality movie. Specifically, assuming exogenous advertising decisions, I estimate that, everything else constant, a 1 per cent increase in advertising spending leads to a 1.32 per cent increase in the movie's market share on average. However, taking the endogeneity of advertising expenditures into account, the same 1 per cent increase in advertising leads to a 1.51 per cent increase in market share.

With my supply and demand estimates at hand, I am able to perform a series of counterfactual exercises. First, I estimate the impact of "coorperative" releases on studio profits, that is, the difference in joint profits between actual releases (equilibrium play) and releases that maximize joint studio profits. I find the average difference to be over $25 million per season. I also show that in the case of maximizing joint studio profits, release dates in a season are more spread out relative to equilibrium. Specifically, in the case of strategic (non-cooperative) play, there is 30 per cent more clashes, i.e. 2 or more blockbuster releases in the same week, compared to cases of maximizing joint studio profits. Finally, I simulate the effect of the recent merger Disney-Fox, and I find a significant increase in profits due to the adjustment of release dates.

There is a large literature in marketing and economics, studying the supply and demand of blockbuster movies. The demand side of my model is similar to Einav (2007), which adopts a discrete choice demand formulation with logit errors. Under this approach, choice sets are assumed to be exogenous. The presence of an endogenous choice set leads to inconsistent estimates due to studios' selection of release dates. Indeed, studios with higher quality movies tend to choose more popular weeks to release these movies. I correct for this selection with a control function-motivated approach introduced in Perrigne, Vuong and Yan (2019). Several authors, including Krider and Weinberg (2008), have looked at the release-date game. However, to the best of my knowledge, mine is the first to do so in a structural empirical approach that includes release-date and advertising decisions, where the theoretical model works for an arbitrary number of studios. Indeed, advertising expenditures are important in the movie industry, and represent on average

30% of production budget. In addition, in an industry where price is essentially constant, advertising expenditures play a similar role as prices in other industries.[2]

The rest of the paper is organized as follows. Section 2 presents an overview of the movie industry and the data. Section 3 surveys the related literature from economics and marketing. Section 4 presents the formal model in a simplified setup for clarity of illustration, and establishes model identification. Section 5 discusses the estimation method and reports estimation results, as well as counterfactual simulations. Section 6 concludes. The Appendix includes the formal model in a generalized setup, some mathematical derivations and hypothesis testing results.

2 The Motion Picture Industry and Data

2.1 The motion picture industry at a glance

Hollywood movies are seen and appreciated by millions of Americans and foreign audience everyday. Movie goers follow Dr. Allen Grant into the wilderness of *Jurassic Park*, immerse themselves in the magical world of *Harry Potter*, and went on a shopping spree for clownfish after the theatrical success of *Finding Nemo*. In addition, Hollywood has produced numerous high-profile movie stars, whose influence go beyond the movie screen. There are also cultural landmark movies like *Jaws* and *Star Wars*, which spawn much excitement and traffic in theaters, as well as discussion and debate at the dinner table.

The impact of Hollywood movies is more than just the incredible experiences for the viewers and the glamour and celebrity appeal. The motion picture industry in the United States is of great economic importance. It has generated over $10 billion inflation adjusted US box office revenue annually, for the past twenty plus years. The industry contributed $100 billion dollars in nominal value added to GDP in 2015. The motion picture industry is also a major private sector employer in nation, contributing to 2.1 million jobs, and

[2] While price has a negative effect on consumer utility, advertising has a positive effect as it positively affects the likelihood of the consumer choosing the movie. Many articles address the issue of price competition in discrete choice models, such as Berry, Levinsohn and Pakes (1995).

$139 billion in total wages in 2016.[3]

2.2 Industry description

The motion picture industry has been a popular topic of research amongst academics, in particular in marketing and more recently economics. Several unique features of the industry make it a fascinating as well as suitable candidate for research: availability of detailed data, short life of products, large outlays of production and advertising, highly variable and difficult-to-predict box office return, rapid decay in box office revenue, the "puzzle" of uniform movie ticket pricing, to name a few. This paper focuses on release timing and advertising: the two major ways studios compete in movie distribution.

The three key stages in the motion pictures value chain are production, distribution and exihibition. Production and distribution are often integrated for the large studios, although the units responsible for the two are usually separate and operate independently. Many independent filmmakers work with larger studios: the former initiate motion pictures and the latter finance their production and manage subsequent theatrical distribution and advertising as discussed by Eliashberg, Elberse and Leenders (2006). The distribution stage involves decisions on release dates, and advertising of movies, which are at the core of this paper. Hereafter, I will refer to distribution studios as "studios". The production stage is largely concerned with the technical making and financing of motion pictures. Exihibition is the screening of the films in theaters. See Eliashberg, Elberse and Leenders (2006) a detailed review of the current practice, critical issues and new directions in production, distribution and exhibition.

The industry is highly integrated, with many studios engaging in or at least (partially) financing the production of movies they distribute. For example, *Mission Impossible - Fallout* (2018) is co-produced by Skydance Productions, Bad Robot and Paramount Pictures, and distributed by Paramount Pictures. If a production company is independent,

[3]Figures provided by The Numbers, the Arts and Cultural Production Satellite Account (ACPSA), and Motion Pictures Association of America (MPAA).

it often has co-production partners with big integrated studios. For example, Skydance Productions often co-produces with Paramount Pictures in the making of many large-budget, action movies such as the *Mission Impossible* movies, *World War Z*, and *Terminator Genisys*.

The process of a movie after its production involves a few key considerations: (i) setting release dates; (ii) setting initial scope of release; (iii) contracting with exhibiting theaters; and (iv) budgeting advertising. The literature mostly focuses on (i) and (iv), related to movie distribution. To name a few, Krider and Weinberg (1998), Radas and Shugan (1998), Eliashberg, Elberse and Leenders (2006), Einav (2010), Belleflamme and Paolini (2018) consider release timing as an important strategic decision for studios. In particular, Einav (2010) shows that opening week revenue accounts for almost 40 per cent of box office revenues on average. Cabral and Natividad (2016) demonstrate that being successful in the opening week is crucial to the overall success of a movie, while Eliashberg, Elberse and Leenders (2006) discuss how a mediocre first week often loses media interest, and exhibitor attention.[4]

Lehmann and Weinberg (2000), Eliashberg, Elberse and Leenders (2006), Elberse and Anand (2007), and Rao et al. (2017) show that advertising is another important strategic variable for studios. Whereas film advertising campaigns are designed far in advance, the actual advertising happens close to the relase date of a movie, peaking immediately before the release date, and may continue for a few weeks after. Given that pricing of a movie ticket has always been quasi-uniform and not dependent on movie budget and other movie characteristics, advertising plays a similar role as prices, in the sense that studios strategically select advertising budgets to compete with other studios in the same season.[5]

[4]Strategic decisions regarding screening, pricing and/or other exhibition-related questions have been studied by De Vany and Walls (1996) (1999), Eliashberg (2005), Weinberg (2005), Davis (2005) (2006a) (2006b) and Orhun et al. (2016). Eliashberg, Elberse and Leenders (2006) and Eliashberg, Weinberg and Hui (2008) also offer comprehensive reviews of related literature.

[5]The determination of admission prices is "puzzle" in view of the uniformity of ticket prices. See Orbach (2004), Weinberg (2005), Orbach and Einav (2007), Courty (2011) and Chung (2015) for some insights on this topic.

Like price, advertising spending affects moviegoer utility, albeit in the opposite direction: higher advertising increases the probability of the consumer choosing the movie.

The Six Players

The domestic motion picture industry is dominated by six studios, commonly known as "The Big Six": Twentieth Century Fox, Paramount Pictures, Sony Pictures Entertainment, Universal Pictures, Warner Bros. Pictures and Walt Disney Studios. Hereafter, we will refer to them as Fox, Paramount, Sony, Universal, Warner and Disney. Together, they make up close to 80% of total grossing in the US. Table 3 shows the top 10 studios by grossings between years 1995-2018. Evident from the figure, there is a significant drop in market shares after the top studios, and all top six studios have market shares exceeding 10 per cent, with no one studio having a significant lead over the other five. The oligopolistic structure of the Big Six studios is consistent throughout the sample period of 1997-2015: the sum of the Big Six studios' annual market shares have always hovered between 66.4 per cent and 84.3 per cent, with the sum being mostly above 80 per cent in recent years, as displayed in Figure 1. There is no clear leader: studios alternate in occupying the spot for the highest yearly market share during 1997-2015. Therefore, the data suggest that the market structure of the top six studios can be described as oligopolistic, and I proceed to model the game as one played between only the Big Six studios.

Each of these six studios has had a long history in Hollywood. The oldest American film studio, Universal, was founded in April 1912. One month later, Paramount Studios was founded. Warner, Disney and Columbia were founded in the 1920s, with Columbia in 1989 becoming part of Sony Pictures. Twentieth Century Pictures was founded in 1933, and in 1935 it merged with Fox Film Corporation (founded in 1915) to create Twentieth Century Fox. In the early days of Hollywood, Metro-Goldwyn-Mayer (MGM) Studios Inc. and Radio-Keith-Orpheum (RKO) Pictures were the other two Golden Age "Big Five" majors, along with Paramount, Warner, and Twentieth Century Fox. MGM operates today as a "mini-major", a term used to describe studios that are markedly smaller than the Big Six but still somewhat non-negligible, whilest RKO ceased production in 1957 and went defunct in 1959. The Big Six structure has been stable over the recent decades. Since 2017, Disney has been engaged in discussion to acquire the majority of Fox. After some hefty upward bid revisions, the deal has closed in March 2019, with a final winning bid of $71 billion.[6] The recent aggressive growth of Disney will probably affect the market structure in the future. In section 5, I simulate the effect of this merger.

2.3 Data

The initial dataset covers all movies released by the Big Six studios between January 1, 1997 and December 31, 2017. For each movie, I collected information on its production budget ($), advertising budget ($), weekly box office revenue ($), weekly number of tickets sold, cumulative domestic box office revenue ($), genre, and distributing studio.[7] Dollar values in the dataset are deflated using the motion picture industry's chain-type price index for industry gross output, an annual time series published by the Bureau of Economic Analysis. Annual population data from U.S. Census and Statistics Canada are used to apportion US/Canada moviegoer population data from the Motion Picture Association of America to arrive at annual US moviegoer population. This is interpolated into weekly moviegoer population figures assuming a linear growth rate. Weekly market shares of each movie is calculated by dividing the number of tickets sold by weekly moviegoer population. Because studios do not always keep track of movie weekly box office revenue in the last few weeks of the movie's life, and box office revenue decays exponentially, I consider the first eight weeks as many previous studies such as Eliashberg and Shugan (1997), Liu (2006), and Einav (2007).

[6] `https://screenrant.com/disney-fox-deal-close-date/`

[7] With the exception of advertising, movie-related data are sourced from *The Numbers*. Total advertising expenditure by movie is sourced from *Ad $ Summary* (for years 1997-2012) or *Ad$pender* (for years 2013-2017), published by Kantar Media.

Seasons

Certain weeks of the year are clearly more popular choices for movie releases than others. Underlying demand for movies is far from constant throughout the year, as people tend to go to the movies more often during the summer (around Memorial Day and July 4) and holiday seasons (Thanksgiving and Christmas). Figure 2 displays the inflation-adjusted sum of opening weekend box office revenues for all movies released by the Big Six studios, from 1997-2015. There are clear peaks in weeks surrounding Memorial Day, July 4, Thanksgiving and Christmas. As opening weekends are crucial for box office success as shown by Eliashberg, Elberse and Leenders (2006) and Cabral and Natividad (2016), Figure 2 sheds important light on the movie release considerations for studios. The observed seasonality in box office revenue results from the combination of fluctuating underlying demand and the endogenous decisions of studios on when to release and the quality of movies released, see Einav (2007). In weeks of high underlying demand, for example around July 4, more people go to the movies, increasing box office revenue. At the same time, studios want to take advantage of this higher demand by releasing more, and larger-budget movies which are aimed for top box office spots, amplifying the flutuations in box office revenue. I am interested in examining the release date and advertising competition amongst the Big Six studios in weeks that have a high underlying demand. These weeks are important for studios' annual performance. I define the seasons in this paper using key dates that are known to have high underlying demand: Memorial Day, July 4, Thanksgiving and Christmas. I define four seasons surrounding these four key dates.[8]

Note that I do not model movie release in a dynamic setting, because big studios are constantly producing new movies that need distribution, and there is ample evidence that studios produced movies assigned for summer release and for Christmas release as

[8]Memorial Day, July 4, Thanksgiving and Christmas typically fall in week 21, 27, 48 and 52 but there are variations from year to year due to calendar week definition. By default, week 1 starts with the week that contains January 1 and all weeks begin on Sunday. So week 1 is only 1 day long if a year starts on a Saturday.

discussed in Krider and Weinberg (1998) and in my own conversations with industry experts. Following Figure 2, there is more than a single peak in box office in the summer and winter seasons. On the other hand, movie production has dramatically increased in the nineties, as displayed in Figure 3. This increasing volume of release has lead to expansions of what is traditionally deemed as summer and winter seasons, consistent with the multiple peaks observed in Figure 2. This justifies my treatment of breaking up the summer and winter months into two separate seasons each: the four seasons in a year surrounding Memorial Day, July 4, Thanksgiving and Christmas represent four independent games for the strategic release date and advertising choices. Regarding the release day of the week, Friday is the norm as shown in Table 5. Thus, the week containing the Friday just before a key date, is a popular week to release a movie. I define two summer seasons of seven weeks each, anchored around Memorial Day and July 4 respectively, and two winter seasons of five weeks each, anchored around Thanksgiving Day and Christmas Day respectively.

Blockbusters

In the motion picture industry, a "blockbuster" typically refers to a movie with a large production budget, and/or large box office. For the purpose of this paper, I opt an ex ante definition, characterizing blockbusters based on the movies' budgets. Since I am interested in the strategic decisions of release dates, it makes sense to use this ex-ante definition of blockbusters, as studios make release decisions prior to knowing box office outcomes, and are likely to pay more attention to the selection of release dates when they have an important stake in the success of the movie. A movie is a blockbuster in my model if it is the top budget movie for a studio in a given season, and its budget is at least 50 per cent of the average budget of the other studios' blockbusters. In the case where a studio does not have a movie a season, the studio is considered not part of the blockbuster release and advertising games for this season. In essence, my definition captures the set of movies that may be viewed as competitors of each other.[9] If there is still none that satisfy this requirement, I consider the studio as choosing not to release a blockbuster in that season. For example, over the Christmas season of 2015, Disney released *Star Wars Ep. VII: The Force Awakens*, with a whopping inflated adjusted budget of \$310,700,000. *The Revenant* by Fox had a budget of \$137,100,000, and Point Break by Warner had a budget of \$106,600,000, which are nonetheless large budgets. In contrast, the highest budget movie by Sony in the same period was *Concussion*, with an inflated-adjusted budget of \$35,500,000, and Paramount's *Daddy's Home* and Universal's Sisters cost \$50,800,000 and \$30,500,000 respectively. While *The Revenant* and *Point Break* may exert competition on *Star Wars*, it is difficult to say the same thing about *Concussion*, *Daddy's Home* and *Sisters*. Therefore, it is best to consider the last three studios as out of the game for this season.

Applying the above definitions of seasons and blockbusters, I end up with 352 blockbuster movies over the years 1997-2015, and 2,816 weekly observations. Table 6 provides descriptive statistics of sample mean, median and standard deviation of budget, first week domestic box office revenue, cumulative domestic revenue, and advertising expenditure. Advertising makes up close to one quarter of total budget (production plus advertising) of an average blockbuster movie, which is around 30 per cent of total production budget. Its correlation with budget is 0.44. It displays significant variance too, with advertising to budget ratio having a coefficient of variation 0.43.

Regarding genres, looking at all movies produced in the last two decades, adventure, action, drama and comedy make up close to 80% of total grossing. Table 4 shows the top 8 genres by grossings between years 1995-2018. There tend to be more adventure and action movies in the summer months, and more family-oriented movies such as *The Grinch*,

[9] In exceptional circumstances where there is an "outlier" blockbuster whose budget significantly exceeds all other movies in the season, I exclude this outlier when computing the average budget of blockbusters. For example, *Dr. Dolittle 2*, *The Fast and the Furious*, *Lara Croft: Tomb Raider*, *Swordfish*, *Pearl Harbor* and *A Knight's Tale* are the top-budget movies in a season in 2001. All have deflated budgets in the range \$63,000,000 -\$89,000,000 except *Pearl Harbor* which has a deflated budget of \$145,000,000. I therefore exclude *Pearl Harbor* and capture all six movies.

The Hobbit, and *The Chronicles of Narnia* around Christmas time. The latter tends to classified as adventure by The Numbers, but since they attract a different demographic, namely younger children and their parents, I redefine such movies as "family". I then aggregate the genres into four categories: Action/Adventure, Family/Comedy, Thriller/-Drama, and Other. Close to 50 per cent of these blockbusters are Action/Adventure, 33 percent are Family/Comedy, 17 per cent are Thriller/Drama and the remaining 2 per cent are Other such as Horror and non-family Musical. I account for the effect of genre by including genre dummy variables in the demand equation.

Apart from the dominance of Action/Adventure movies, another notable trait 202 out of 352 blockbusters share is that they come from something "well-known": 105 blockbusters are sequel or prequel to a previous well-known film, 33 blockbusters are remakes/reboots of a previous well-known film, and 116 are based off a best-selling book or a popular musical.[10] I define a dummy variable "fame" which takes value 1 if a movie satisfies one of the above, and 0 otherwise.

3 Related Literature

The past thirty years have witnessed an abundant literature on the motion picture industry, advertising, demand estimation and entry games. This literature overlaps marketing, economics and management science. Hereafter, I review this literature in two topics: demand estimation and entry games. In each subsection, I also discuss papers on the movie industry.

Demand Estimation

The seminal paper BLP (1995) estimates demand for differentiated products via a discrete choice model, in parallel of Bertrand competition for prices, taking the product choice set as given, and exogenous product characteristics. Some recent developments include

[10]Some blockbusters possess more than one of these features, hence the sum of 105, 33 and 116 exceeds 202.

Ackerberg, Crawford and Hahn (2011) and Fan (2013), where BLP is extended to endogenous product characteristics. Li et al (2018), Ciliberto et al (2018) and PVY (2019) build upon the BLP framework and adds firm strategic entry to the model, thereby endogenizing the product choice set. Lastly, Compiani (2019) develops a non-parametric estimator of demand.

The treatment of advertising in discrete choice models is either exogenous, with advertising entering as a covariate, or endogenous when the firms choose advertising to optimize various objectives. In the former line of research, Ackerberg (2001) distinguishes informative and prestige effects of advertising, while Dubé, Hitsch and Manchanda (2005) study the long-run effects of advertising. In the latter line of research, Goeree (2008) allows advertising to influence the product choice set, and Murry (2017) models firms choosing advertising to maximize profit, with advertising entering the consumer's indirect utility. Shum (2004) considers brand loyalty in which advertising, and its interaction terms with price and/or past use of the product, enter the consumer's indirect utility. Luan and Sudhir (2010) forecast advertising responsiveness in the U.S. DVD market, correcting for "slope-endogeneity" of advertising[11] with a control function approach.

Regarding the empirical analysis of demand in the movie industry, a reduced form literature with Eliasberg and Shugan (1997), Eliashberg et al. (2000), Elberse (2007), Eliashberg et al. (2014), Packard et al. (2016) study the impact of critics' reviews, star actors, networks of cast and crew, and opening weekend box office on overall revenue performance. Lehmann and Weinberg (2000), Elberse and Anand (2007) and Rao et al. (2017) consider the impact of advertising on box office performance. Elberse and Anand (2007) found on average a negative return on advertising, while the return varies greatly across movies of different qualities as measured by critics' reviews. Fewer papers have used a discrete choice structural approach. Einav (2007) focuses on the seasonality in demand, while Moul (2007) assesses the impact of word-of-mouth advertising.

[11] This refers to where the responsiveness of demand to advertising is correlated with firms' private information.

Estimation of Entry/Release Timing

Starting with Berry (1992) and Bresnahan and Reiss (1990, 1991), there is an important literature modeling entry as the outcome of an equilibrium under complete information. In an incomplete information setting, Seim (2006) considers an entry model with endogenous product-type choice. Sweeting (2009) develops an entry game to study strategic timing. In all these papers, the firms' payoffs are reduced-form, and in the last two papers, the firms' private information is assumed to be mutually independent. In contrast, my paper allows firms' private information to be correlated, and incorporate demand in the firms' payoff.

Regarding the literature on movie release timing, Radas and Shugan (1998) and Krider and Weinberg (2008) study how seasonality affects movie release timing. Belleflamme and Paolini (2018) develop a model in which firms select a release date after investing in attractiveness of their movie. Einav (2010) considers a sequetial move timing game, and examines the extent to which distributors mis-estimate underlying demand, leading to suboptimal release dates. Lehmann and Weinberg (2000) find that studios wait too long to release movies to video, while Natividad (2013) studies how unexpectedly successful films affect future distribution. Jones and Ritz (1991) and Elberse and Eliashberg (2003) study demand and supply dynamics present in industries where products are sequentially released. Like other entry papers discussed above, payoffs of studios are reduced-form.

4 The Model

In view of the above discussion and evidence, I develop a two-stage game of release timing and subsequent advertising, with a discrete choice model for movie demand. Introducing the demand leads to structural firm payoffs in the release game, while the release timing endogenizes the consumer's choice set in a given week in the demand model. I characterize the equilibrium advertising and release strategies. Both release and advertising are under incomplete information, though when choosing advertising, studios update their beliefs about rivals' movie qualities, having observed their release date choices.

Discrete Choice Demand Model

I follow the literature on discrete choice models with logit demand. In each week, moviegoers derive utilities from each alternative/blockbuster movie in a set of possible alternatives $C \subseteq \{1, 2, ..., J\}$, $J \geq 1$, or the outside option, "0". The outside option includes non-blockbuster movies and other entertaining options. Studios are single-product, therefore $j \in C$ can be interpreted equivalently as studio j or blockbuster j, and J is the number of blockbusters in a season.

Each calendar week is denoted by $t = 1, ..., 52$. Let W denote the length, in weeks, of a release season, $W = 4$ in winter and $W = 7$ in summer. The week of release in a season chosen by studio j is denoted R_j, $R_j \in \{1, ..., W\}$, while $\tilde{R}_j$ the corresponding calendar week. For instance, if a season is four weeks long, and consists of calendar weeks 25 to 28, then release by j in week 2 is denoted $R_j = 2$, and $\tilde{R}_j = 26$. Thus, the number of weeks lapsed since the initial release of a blockbuster movie, is $t - \tilde{R}$. I follow Einav (2007), where the utility is a linear function of $t - \tilde{R}$. Under Type I Extreme Value-distributed errors, the linearity of utility in $t - \tilde{R}$ leads to an exponential decay in the market share, which is consistent with the observed evolution of weekly box office revenue. This exponential decay potentially explains the importance of opening week performance.

Each blockbuster movie $j \in \{1, ..., J\}$ is viewed as a bundle of observed characteristics $X_j \in \mathbb{R}^{\dim X}$ and unobserved characteristics $\xi_j \in \mathbb{R}$. The latter is commonly referred to as the "product quality", and is studio j's private information prior to release. More specifically, ξ_j is unobserved to rival studios at the time when studios make release and advertising decisions. Studio j knows their competitors' movie characteristics X_{-j}. I allow the possibility of the ξ_js being correlated across j, with joint distribution $F_{\xi_1,...,\xi_J}(\cdot, ..., \cdot)$. The quality ξ_j consitutes the source of incomplete information in the studios' release decision. On the other hand, once movie j is released, consumers become aware of the movie quality ξ_j through critics' reviews, social media etc.

Advertising expenditure on movie j is denoted $A_j \in \mathbb{R}^+$. The higher the A_j, the more likely that consumer will choose to see movie j. Therefore, including A_j in the utility of the consumer choosing movie j reflects that higher A_j increases the probability of choosing movie j.

The utility of the consumer i for choosing blockbuster j in week t is:

$$u_{ijt} = X_j'\beta - \lambda(t - \tilde{R}_j) + \alpha \log(A_j) + \xi_j + \eta_{jt} + \epsilon_{ijt},$$

where $\eta_{jt} \in \mathbb{R}$ is a mean-zero movie-week independent shock that is unknown to every firm prior to release, distributed as $F_{\eta_j}(\cdot)$, $j = 1, ..., J$, and $\epsilon_{ijt} \in \mathbb{R}$ is an indiosyncratic demand shock distributed as Type I Extreme Value.

It is standard in the discrete choice literature to normalize the utility from consuming the outside option to $u_{i0t} = \epsilon_{i0t}$. In the context of movie demand, it is important to acknowledge that the demand seasonality is a combination of the underlying seasonality, which captures preference for some weeks such as holiday weeks over others, and seasonality due to the number and quality of movies available. We can expect that in weeks with a strong underlying demand such as around Memorial Day and Christmas, more and potentially better quality movies are on display. Einav (2007) disentangles these two effects by including a week dummy τ_t in the outside utility, thus the outside utility becomes

$$u_{i0t} = -\tau_t + \epsilon_{i0t},$$

where ϵ_{i0t} is also distributed as Type I Extreme Value.
The market share of movie j in week t, s_{jt}, is expressed as

$$s_{jt} = \begin{cases} 0 & \text{if } j \text{ is not present in week } t \\ \dfrac{\exp^{X_j'\beta - \lambda(t-\tilde{R}_j) + \alpha \log(A_j) + \xi_j + \eta_{jt}}}{\exp^{-\tau_t} + \sum_{j' \in C_t} \exp^{X_{j'}'\beta - \lambda(t-\tilde{R}_j) + \alpha \log(A_{j'}) + \xi_{j'} + \eta_{j't}}} & \text{otherwise,} \end{cases}$$

where C_t denotes the set of available blockbusters in week t.

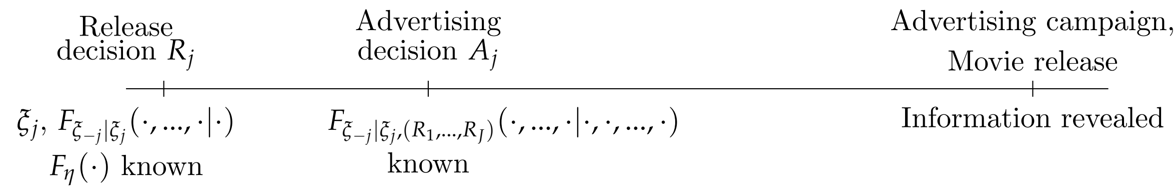
Taking ratios of s_{jt} and s_{0t} and applying the natural logarithm gives the logit demand equation

$$\log\left(\frac{s_{jt}}{s_{0t}}\right) = X_j'\beta + \tau_t - \lambda(t - \tilde{R}_j) + \alpha \log(A_j) + \xi_j + \eta_{jt}. \tag{1}$$

As discussed previously, advertising is related to a blockbuster's quality, ξ_j. Having A_j and ξ_j in the utility function means that movie j's market share will depend on both of these variables. Therefore, studio j's choice of A_j will necessarily be a function of ξ_j, making advertising an endogenous variable in the demand equation.

Firms' Strategic Decisions on Release Date and Advertising

The firms' model is motivated by data and facts about the motion picture industry. As discussed in the Data section, six major players dominate the industry in terms of market share. Their release decisions are made in advance, and they use advertising to influence consumer choice and drive moviegoers to the theaters, especially on crucial opening weekends. I illustrate the timing of the model, as well as the corresponding information structures in the following figure. I then discuss the first and second stages of the model in details.



First Stage: Release Date Decision

Consider a studio with a blockbuster movie that is near or at its production completion, and needs to decide on a week to release it next summer. It has seen its own movie, or at least knows enough about it, through excerpts of the filming, meetings with the crew and the cast, on-site visits, and so on, to assess the movie quality. But it does not know as much about its rivals' movies beyond what is public information such as title, cast, production company and potentially budget. This motivates the First Stage of the game to be characterized by incomplete information in the "movie quality", ξ_j. Since $\xi_1, ..., \xi_J$ are potentially dependent, studios know the joint distribution $F_{\xi_1,...,\xi_J}(\cdot, ..., \cdot)$. Conditioning on j's own ξ_j may provide studio j with more information about $\xi_{-j} \equiv \{\xi_{j'} : j' \in \{1, ..., J\} \backslash j\}$, leading to the belief distribution $F_{\xi_{-j}|\xi_j}(\cdot|\cdot)$. Once a movie is released or near its release date, the movie quality becomes known to everyone via critics' reviews, social media etc., hence its presence in the moviegoer's utility function. For the rest of the paper, I will refer to this studio private information prior to release as "quality". In the industrial organization literature, $\mathbb{E}(\xi_j)$ is commonly referred to as the "brand effect" of product j.

Each studio chooses optimally the release week of their blockbuster movie. As discussed above, studio j knows ξ_j, but not ξ_{-j}. No studio knows the movie-week shocks, η_{jt}, $j = 1, ..., J$, $t = 1, ..., W$, but their distributions $F_{\eta_1}(\cdot)$,...,$F_{\eta_J}(\cdot)$ are known. As usual, studios also know the parameters of demand, and the observable characteristics of each studio's blockbuster, X_j. The studios' release choices $R_1, ..., R_J$ characterize the weekly sets of available blockbusters in a season, C_t, $t = 1, ..., W$. The consumer's weekly choice set is therefore $\{0\} \cup C_t$, $t = 1, ..., W$.

Second Stage: Advertising Decision

Once release dates are announced, studios decide how much to spend on advertising. Actual advertising campaign happens close to movie release, its budgeting however, happens much earlier, as studios need to secure slots in various channels of advertising, as pointed out by Elberse and Anand (2007). In view of this, I still assume ξ_{-j} remains unknown to studio j, $j = 1, ..., J$ when they make advertising decisions.

As mentioned previously, advertising plays the role of price in a standard BLP model. Studios play a Bertrand-style competition game on advertising, choosing their optimal advertising spending to best respond to any arbitrary advertising levels chosen by their rivals, given blockbuster observable characteristics X_j $\forall j$, its own movie quality ξ_j, the joint quality distribution $F_{\xi_1,...,\xi_J}$ $(\cdot, ..., \cdot)$, and movie week shock distribution F_{η_j} $\forall j$. The studios also observe everyone's release date decisions R_j $\forall j$, the "market structure" of the season, which maps to weekly choice sets moviegoers face. This leads to a new belief distribution $F_{\xi_{-j}|\xi_j,(R_1,...,R_J)}$ $(\cdot, ..., \cdot|\cdot, \cdot, ..., \cdot)$.

To focus on the main ideas of the model, and to avoid cumbersome notation in the text, I formally present after a model with 2 studios releasing over a season of 2 weeks, that is $J = 2$, $W = 2$. The case for an arbitrary J studios and W weeks extends from here, and is presented in the Appendix.

The Model for 2 studios and 2 week choices

For each $j \in \{1,2\}$, the week of release in a given season, R_j, is chosen from the set $\{1,2\}$. Let M_t denote the product of moviegoer population in week t and ticket price in week t. That is, M_t is the market size (in dollar terms) in week t. The Data section above presents a detailed account on how I obtain an accurate measure of market size. I follow Eliashberg and Shugan (1997), Liu (2006) and Einav (2007), and assume each blockbuster movie screens for eight weeks. Let $\pi_{j,(R_1,R_2)}$ denote the profit of studio j when the release decision is (R_1, R_2). For each of the four possibilities of release outcomes: $(R_1, R_2) = (1,1), (1,2), (2,1)$ or $(2,2)$, I derive the corresponding profit:[12]

Case $(R_1, R_2) = (2,2)$: Both studios release in the second week.

$\pi_{j,(2,2)} = M_2 s_{j2} + M_3 s_{j3} + ... + M_9 s_{j9} - A_j = \sum_{t=2}^{9} M_t s_{jt} - A_j$ for both studios.

Case $(R_1, R_2) = (1,2)$: Studio 1 (resp. 2) releases in the first (resp. second) week.

$\pi_{1,(1,2)} = M_1 s_{11} + M_2 s_{12} + ... + M_8 s_{18} - A_1 = \sum_{t=1}^{8} M_t s_{1t} - A_1$,

$\pi_{2,(1,2)} = M_2 s_{22} + M_3 s_{23} + ... + M_9 s_{29} - A_2 = \sum_{t=2}^{9} M_t s_{2t} - A_2$.

Note that in Case $(R_1, R_2) = (1,2)$, s_{11} and s_{29} are monopoly shares and the rest are duopoly shares.

Case $(R_1, R_2) = (2,1)$ is similar to the case $(R_1, R_2) = (1,2)$, with the studio indices switched places.

Case $(R_1, R_2) = (1,1)$: Both studios release in the first week.

$\pi_{j,(1,1)} = M_1 s_{j1} + M_2 s_{j2} + ... + M_8 s_{j8} - A_j = \sum_{t=1}^{8} M_t s_{jt} - A_j$ for both studios.

Let m and d superscripts denote the monopoly and duopoly cases, respectively. To summarize, for any calendar week release $(\tilde{R}_1, \tilde{R}_2)$ corresponding to (R_1, R_2), the profit of studio j is

$\pi_{j,(\tilde{R}_1,\tilde{R}_2)} = \sum_{t=\tilde{R}_j}^{\tilde{R}_j+7} M_t s_{jt} - A_j$, where the market shares are defined as

[12] Movie budget, i.e. production cost, is sunk. Therefore I do not include it in the profit for the analysis of release and advertising.

$$s_{jt} = \begin{cases} s^m_{jt} & \text{if } \{\tilde{R}_j < \tilde{R}_{j'}, t < \tilde{R}_{j'}\} \text{ or } \{\tilde{R}_j \geq \tilde{R}_{j'}, t > \tilde{R}_{j'} + 7\} \\ s^d_{jt} & \text{if } \{\tilde{R}_j < \tilde{R}_{j'}, t \geq \tilde{R}_{j'}\} \text{ or } \{\tilde{R}_j \geq \tilde{R}_{j'}, t \leq \tilde{R}_{j'} + 7\}, \end{cases}$$

where $\tilde{R}$ refers to the calendar week equivalent of R_j. The market share can be expressed as:

$$s_{jt} = s^m_{jt}[\mathbb{1}(\tilde{R}_j < \tilde{R}_{j'})\mathbb{1}(t < \tilde{R}_{j'}) + \mathbb{1}(\tilde{R}_j \geq \tilde{R}_{j'})\mathbb{1}(t > \tilde{R}_{j'} + 7)]$$

$$+ s^d_{jt}[\mathbb{1}(\tilde{R}_j < \tilde{R}_{j'})\mathbb{1}(t \geq \tilde{R}_{j'}) + \mathbb{1}(\tilde{R}_j \geq \tilde{R}_{j'})\mathbb{1}(t \leq \tilde{R}_{j'} + 7)].$$

From the demand logit model, we know that the monopoly and duopoly market shares take the following expressions respectively, for $j = 1, 2$:

$$\begin{aligned} s^m_{jt} &= \frac{\exp^{X'_j\beta - \lambda(t-\tilde{R}_j) + \alpha\log(A_j) + \xi_j + \eta_{jt}}}{\exp^{-\tau_t} + \exp^{X'_j\beta - \lambda(t-\tilde{R}_j) + \alpha\log(A_j) + \xi_j + \eta_{jt}}} \\ &= s^m_{jt}(X_j, \tilde{R}_j, A_j, \xi_j, \eta_{jt}), \end{aligned}$$

$$\begin{aligned} s^d_{jt} &= \frac{\exp^{X'_j\beta - \lambda(t-\tilde{R}_j) + \alpha\log(A_j) + \xi_j + \eta_{jt}}}{\exp^{-\tau_t} + \exp^{X'_j\beta - \lambda(t-\tilde{R}_j) + \alpha\log(A_j) + \xi_j + \eta_{jt}} + \exp^{X'_{j'}\beta - \lambda(t-\tilde{R}_{j'}) + \alpha\log(A_{j'}) + \xi_{j'} + \eta_{j't}}} \\ &= s^d_{jt}(X_j, X_{j'}, \tilde{R}_j, \tilde{R}_{j'}, A_j, A_{j'}, \xi_j, \xi_{j'}, \eta_{jt}, \eta_{j't}). \end{aligned}$$

To analyze the firms' decisions, I start by looking at the second stage of the game.

2nd Stage: Advertising Decision

Given the information structure for the second stage as discussed previously, firm j chooses A_j given the information $(X_1, X_2, \tilde{R}_1, \tilde{R}_2, \xi_j)$, and hence the belief distribution $F_{\xi_{-j}|\xi_j,(\tilde{R}_1,\tilde{R}_2)}(\cdot|\cdot,\cdot,\cdot)$. As before, movie-week shocks are unknown but their distributions $F_{\eta_1}(\cdot)$ and $F_{\eta_2}(\cdot)$ are known. Thus, studio j solves the following profit maximization problem under incomplete information:

$$\max_{A_j} \mathbb{E}[\pi_{j,(\tilde{R}_1,\tilde{R}_2)}(X_1, X_2, A_1, A_2, \xi_2, \xi_2, \eta_1, \eta_2|\xi_j, R_1, R_2)],$$

where the expectation is taken with respect to ξ_{-j}, η_1 and η_2. This leads to the first-order condition

$$\frac{\partial \mathbb{E}[\pi_{j,(R_1,R_2)}(X_1, X_2, A_1, A_2, \xi_2, \xi_2, \eta_1, \eta_2)|\xi_j, \tilde{R}_1, \tilde{R}_2]}{\partial A_j} = 0, \;\; j = 1, 2.$$

Solving for (A_1, A_2) simultaneously from the above two first-order conditions yields the optimal advertising strategies:

$$A_j^* = A^*_{j,(\tilde{R}_1,\tilde{R}_2)}(X_1, X_2, \xi_j), \; j = 1, 2, \tag{2}$$

and the corresponding maximized profits:

$$\pi^*_{j,(\tilde{R}_1,\tilde{R}_2)}(A_1^*, A_2^*) = \pi^*_{j,(\tilde{R}_1,\tilde{R}_2)}(X_1, X_2, \xi_1, \xi_2), \; j = 1, 2.$$

1st Stage: Release Decision Under Incomplete Information

In this stage, each studio choose a release week for their blockbuster movie. The desirabilities of the weeks in a season is ranked by their potential underlying demand. Suppose the season consists of calendar weeks 48 and 49, and that Thanksgiving Day, the week of higher underlying demand, falls in week 48. In this case, week 1 is the most desirable. Consider another 2-week season of weeks 51 and 52. Christmas Day falls in week 52, so week 2 in the season is the most desirable.

To account for this, I rank weeks 1 and 2 in the season according to the week dummy $\tau_t, t = 1, 2$ and define $\tau^{-1}[1]$ as period with the highest τ_t, and $\tau^{-1}[2]$ as period with the second highest τ_t.[13] Therefore, a firm's choice R_j is a choice of $\tau^{-1}[w]$, $w = 1, ..., W$. Here, $W = 2$.

Recall the information structure in the first stage: firm j knows ξ_j but not ξ_{-j}. Neither firm knows their corresponding movie-week shock η_j. A studio chooses the release week that maximizes its expected payoff given the release strategies of other firms. This is in the spirit of Sweeting (2009), who models the strategic timing of radio commercials under incomplete information, but with reduced-form payoffs.

Studio j chooses $\tau^{-1}[w]$ if and only if its expected payoff under $\tau^{-1}[w]$ exceeds its expected payoff under $\{1, 2\}\backslash\tau^{-1}[w]$, $w = 1, 2$. I consider monotone Bayesian Nash equilibrium strategies. The higher the quality of its blockbuster movie, the more "aggressive" a studio is likely to be in setting its release date, in the sense that it is likely going to select a week with high underlying

[13]Note that in estimation, since we have multiple seasons in a year, we will index τ by t where t denotes the calendar week as opposed to the t^{th} week in the game.

demand. A studio with a low quality movie may optimally choose the weaker week, if it expects the rival to be stronger, and thus likely to release in the stronger week.

Let ξ_j^*, $j = 1,2$ denote the "cut-off", or "threshold" qualities of studios 1 and 2 respectively under the monotone Bayesian Nash equilibrium. When ξ_j, $j = 1,2$ is at the threshold level, studio j is indifferent between releasing in the first or second week. Thus the release week strategy of studio j, $j = 1,2$, is a mapping $r(\cdot)$ from ξ_j to a week choice as

$$R_j = r(\xi_j) = \begin{cases} \tau^{-1}[1] & \text{if } \xi_j \geq \xi_j^* \\ \tau^{-1}[2] & \text{if } \xi_j < \xi_j^* \end{cases} \qquad j = 1,2,$$

where the equilibrium thresholds ξ_j^*, $j = 1,2$ solve simultaneously studio 1 and studio 2's indifferent conditions:

$$\mathbb{E}\left[\pi^*_{1,(\tau^{-1}[1],R_2)}(\xi_1^*,\xi_2)|\xi_1^*\right] = \mathbb{E}\left[\pi^*_{1,(\tau^{-1}[2],R_2)}(\xi_1^*,\xi_2)|\xi_1^*\right] \text{, and}$$

$$\mathbb{E}\left[\pi^*_{2,(R_1,\tau^{-1}[1])}(\xi_1,\xi_2^*)|\xi_2^*\right] = \mathbb{E}\left[\pi^*_{2,(R_1,\tau^{-1}[2])}(\xi_1,\xi_2^*)|\xi_2^*\right],$$

where the expectations are taken with respect to ξ_{-j}, R_{-j}, η_1 and η_2 conditioning on firm j's private information ξ_j, $j = 1,2$. Recall η_j is independent of ξ_j, $j = 1,2$, and studios know the joint distribution $F_{\xi_1,\xi_2}(\cdot,\cdot)$, and the distributions of η_{jt}, $F_{\eta_j}(\cdot)$, $j = 1,2$. That is, the Bayesian Nash equilibrium tells us that when a studio has movie quality above (resp. below) its threshold, it will choose the stronger (resp. weaker) week.

Explicitly expressing the expectations in say studio 1's indifference condition leads to:

$$\begin{aligned}
&F_{\xi_2|\xi_1}(\xi_2^*|\xi_1^*)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\xi_2^*}\pi^*_{1,(\tau^{-1}[1],\tau^{-1}[2])}(\xi_1^*,\xi_2)dF_{\xi_2|\xi_1}(\xi_2|\xi_1^*)dF_{\eta_1}(\eta_1)dF_{\eta_2}(\eta_2)\\
&\quad+(1-F_{\xi_2|\xi_1}(\xi_2^*|\xi_1^*))\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{\xi_2^*}^{\infty}\pi^*_{1,(\tau^{-1}[1],\tau^{-1}[1])}(\xi_1^*,\xi_2)dF_{\xi_2|\xi_1}(\xi_2|\xi_1^*)dF_{\eta_1}(\eta_1)dF_{\eta_2}(\eta_2)\\
&=F_{\xi_2|\xi_1}(\xi_2^*|\xi_1^*)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\xi_2^*}\pi^*_{1,(\tau^{-1}[2],\tau^{-1}[2])}(\xi_1^*,\xi_2)dF_{\xi_2|\xi_1}(\xi_2|\xi_1^*)dF_{\eta_1}(\eta_1)dF_{\eta_2}(\eta_2)\\
&\quad+(1-F_{\xi_2|\xi_1}(\xi_2^*|\xi_1^*))\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{\xi_2^*}^{\infty}\pi^*_{1,(\tau^{-1}[2],\tau^{-1}[1])}(\xi_1^*,\xi_2)dF_{\xi_2|\xi_1}(\xi_2|\xi_1^*)dF_{\eta_1}(\eta_1)dF_{\eta_2}(\eta_2),
\end{aligned}$$

where, since the week choice is discrete, the probabilities of studio 2 releasing in the stronger week $\tau^{-1}[1]$, and studio 2 releasing in the weaker week $\tau^{-1}[2]$, given ξ_1^*, are $(1 - F_{\xi_2|\xi_1}(\xi_2^*|\xi_1^*))$ and

$F_{\xi_2|\xi_1}(\xi_2^*|\xi_1^*)$ respectively. The first two integrals are taken with respect to η_1 and η_2 respectively, and the third integral is taken with respect to ξ_2 conditional on $\xi_1 = \xi_1^*$.

Similarly, studio 2's indifference condition can be expressed as:

$$F_{\xi_1|\xi_2}(\xi_1^*|\xi_2^*) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\xi_1^*} \pi^*_{2,(\tau^{-1}[2],\tau^{-1}[1])}(\xi_1, \xi_2^*) dF_{\xi_1|\xi_2}(\xi_1|\xi_2^*) dF_{\eta_1}(\eta_1) dF_{\eta_2}(\eta_2)$$
$$+(1 - F_{\xi_1|\xi_2}(\xi_1^*|\xi_2^*)) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{\xi_1^*}^{\infty} \pi^*_{2,(\tau^{-1}[1],\tau^{-1}[1])}(\xi_1, \xi_2^*) dF_{\xi_1|\xi_2}(\xi_1|\xi_2^*) dF_{\eta_1}(\eta_1) dF_{\eta_2}(\eta_2)$$
$$= F_{\xi_1|\xi_2}(\xi_1^*|\xi_2^*) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\xi_1^*} \pi^*_{2,(\tau^{-1}[2],\tau^{-1}[2])}(\xi_1, \xi_2^*) dF_{\xi_1|\xi_2}(\xi_1|\xi_2^*) dF_{\eta_1}(\eta_1) dF_{\eta_2}(\eta_2)$$
$$+(1 - F_{\xi_1|\xi_2}(\xi_1^*|\xi_2^*)) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{\xi_1^*}^{\infty} \pi^*_{2,(\tau^{-1}[1],\tau^{-1}[2])}(\xi_1, \xi_2^*) dF_{\xi_1|\xi_2}(\xi_1|\xi_2^*) dF_{\eta_1}(\eta_1) dF_{\eta_2}(\eta_2)$$

The same comments apply as for studio 1's difference condition.

Existence of Bayes Nash Equilibrium

The entry game I consider is static, and falls under finite-action games, where the set of actions is compact. I assume that the joint density of ξ_j, $j = 1, ..., J$, is bounded and atomless. Furthermore, the profit function of studio j, $j = 1, ..., J$ is continuous, finite and integrable with respect to ξ_j. Under reasonable distribution of $F_{\xi_1,...,\xi_J}(\cdot, ..., \cdot)$, the entry game also satisfies the single crossing property as defined in Athey (2001). Specifically, ξ_j enters positively in the consumer's utility function, i.e. higher ξ_j means the consumer is more likely to choose blockbuster j in week t, leading to a higher market share for blockbuster j in week t. Therefore ξ_j is how inherently "good" blockbuster j is, thus if studio j prefers a week with higher underlying demand to one with a lower underlying demand when $\xi_j = \xi_L$, then the week with higher underlying demand should remain preferable to the one with lower underlying demand, when $\xi_j = \xi_H$, $\xi_H > \xi_L$, as long as studio j's beliefs about ξ_{-j} are "reasonable".[14] The existence of a monotone Bayes Nash equilibrium in my entry model then follows Theorem 1 in Athey (2001).

5 Model Identification and Estimation

This section addresses the identification of model parameters developed above. Each release season as defined in the data section is an independent replication of the model. I first discuss the selection issue arising from ξ_j as a result of the endogenous choice set in the demand model.

[14]Distributions $F_{\xi_1,...,\xi_J}(\cdot, ..., \cdot)$ such that when firm j draws a higher ξ_j, it becomes more "pessimistic" and believes its rivals' ξ_{-j} may be even higher, so that it chooses a more "conservative" release week with lower underlying demand, are excluded. Indeed, I estimate the empirical pdfs of ξ_js after model estimation, and they appear mostly Normal.

I then provide sufficient conditions for identification. After, I develop a two-step estimation procedure, accounting for the selection issue and the endogeneity of advertising.

Endogenous choice set and selection

Standard estimation of the demand equation (1) will lead to inconsistent estimates, due to the issue of selection. Specifically, the mean independence restriction necessary for consistency, i.e. $\mathbb{E}(\xi_j|X_1, ..., X_J, R_1, ..., R_J)$ equals a constant, is not satisfied. To simplify notations, let $\mathbf{X} \equiv (X_1, ..., X_J)$ and $\mathbf{R} \equiv (R_1, ..., R_J)$. Since the realization of ξ_j affects studio j's release choice, for instance, a studio that draw higher ξ_j is more likely to release in a week with higher underlying demand. Therefore, the expectation of ξ_j is likely to be higher when studio j releases in a high week.

To account for this selection bias on ξ_j, I adopt a control function approach as in PVY (2019). The idea behind the control function is that it captures the conditional mean $\mathbb{E}(\xi_j|\mathbf{X}, \mathbf{R})$, therefore controlling for the problem of endogenenous choice set (see Hausman (1978), Heckman (1979), Rivers and Vuong (1988), Robinson (1988), Das, Newey and Vella (2003) and PVY (2019) for a control function approach in model estimation with selection). Denote the probability that studio j releases in week $\tau^{-1}[w]$, or "entry probability" as $p_{j\tau^{-1}[w]}(\mathbf{X})$, and the probability of observing release pattern $\mathbf{R}$ as $p(\mathbf{R}|\mathbf{X})$. I will refer to the latter as the "probability of market structure $\mathbf{R}$". Note that the former is a function of all studios' observable characteristics, $\mathbf{X}$. Let $p(\mathbf{X})$ be the vector containing these probabilities,

$$p(\mathbf{X}) \equiv \left[[p_{j\tau^{-1}[w]}(\mathbf{X})]'_{j=1,...,J} \; p(\mathbf{R}|\mathbf{X}) \right].$$

Assume that any market structure can occur with some positive probability, i.e. $\Pr(\mathbf{R}|\mathbf{X}) \geq 0$ for every $\mathbf{R}$, where $R_j \in 1, ..., W$ for each j. Furthermore, assume that the unobserved qualities $\xi \equiv (\xi_1, ..., \xi_J)$ are independent of $\mathbf{X}$. That is,

$$\xi \perp \mathbf{X}$$

PVY (2019) extend the argument in Das, Newey and Vella (2003) to show that

$$\mathbb{E}[\xi_j|\mathbf{X}, \mathbf{R}] = \lambda^{\xi}_{j:\mathbf{R}}[p(\mathbf{X})] \tag{3}$$

for some control function $\lambda^{\xi}_{j,\mathbf{R}}(\cdot)$. Therefore, firm j's control function depends not only on j's entry probability, but every other firm's entry probabilities and the probability of the market structure as well.

Identification of model parameters

The model parameters presented in the previous section to be identified are $\theta = (\beta', \tau', \lambda, \alpha)'$.

First of all, advertising A_j is endogenous, and needs to be corrected with an instrumental variable. I use budget B_j as its instrument. The identifying assumptions for a valid instrument are: (i) B_j is correlated with A_j, and (ii) B_j is uncorrelated with $\xi_j + \eta_{jt}$, the error term in the demand equation. Assumption (i) is easily satisfied, given that higher budget movies also tend to spend more on advertising, see empirical evidence in the Data section, where the correlation between B_j and A_j is 0.44. The endogeneity of A_j stems from its correlation with ξ_j, as emphasized in (2). Namely, at the time studios make decisions regarding advertising, they have some private information, ξ_j, that is unknown to other studios. Therefore, given the importance of opening weekend performance on overall box office, studios with lower ξ_j may advertise more aggressively. Regarding (ii), I need to justify that B_j, on the other hand, is uncorrelated with ξ_j. Higher budgets are often related to heavy use of special effects, 3D animation and more generally action/adventure movies, which do not reflect the quality of the movie. Furthermore, B_j does not enter consumers' utility specification, as budget is unlikely to affect consumers' movie choice beyond information that is available in X_j.

I then proceed with an Instrumental Variable (IV) based estimator. In the first stage, $\tilde{A}_j \equiv \mathbb{E}[A_j|X_j, B_j]$ is identified from the regression of A_j on the instrument B_j and exogenous regressors X_j. Let V_j be the residual from the first stage regression. Coupled with equation (3), it leads to the partial linear regression

$$\log\left(\frac{s_{jt}}{s_{0t}}\right) = X_j'\beta + \tau_t - \lambda(t - \tilde{R}_j) + \alpha \log \tilde{A}_j + \lambda^{\xi}_{j:\mathbf{R}}[p(\mathbf{X})] + \varphi_{j:\mathbf{R}} \tag{4}$$

for some composite error $\varphi_{j:R_1,\ldots,R_6} = (\xi_j - \mathbb{E}[\xi_j|\mathbf{X}, \mathbf{R}]) - \alpha V_j$, which by construction satisfies $\mathbb{E}[\varphi_{j:\tilde{R}}|\mathbf{X}, \mathbf{R}] = 0$. By the Law of Iterated Expectations and noting that $(X_j, \tilde{A}_j, p_{j\tau^{-1}[w]}(\mathbf{X}))$ are functions of $\mathbf{X}$, $\mathbb{E}[\varphi_{j:\tilde{R}}|\mathbf{X}, \tilde{A}_j, p_{j\tau^{-1}[w]}(\boldsymbol{X}), \mathbf{R}] = 0$. Essentially, incorporating the control function $\lambda^{\xi}_{j:\mathbf{R}}(\cdot)$ enables me to re-write the demand equation as one that has a zero-mean error term conditional on observables and choice sets.

As discussed above, $\tilde{A}_j$ is identified. The vector $p_{j\tau^{-1}[w]}(X)$ is identified because each component is identified from studio j's release choice, given X, or the market structure $\mathbf{R}$ given $\mathbf{X}$. Therefore, the regressors in equation (4) are identified. Following Robinson (1998), parameters $\theta = (\beta', \tau', \lambda, \alpha)'$ in the partially linear regression depicted by equation (4) are identified.

A Two Step Estimation Procedure

The estimation procedure closely follows the identification argument. I employ an instrumental variable approach to correct for the endogeneity of advertising. The endogenous choice set is corrected by the control function $\lambda^{\xi}_{j:\mathbf{R}}[p(\mathbf{X})]$. In estimation, the vector of probabilities $p(\mathbf{X})$ is replaced by $\hat{p}(\mathbf{X})$, which is estimated from data.

My estimation procedure is therefore in two-step. In the first step, I estimate $p(\mathbf{X})$ from observed release choices of studios. In the second step, I estimate the demand equation (4), with $p(\mathbf{X})$ in the control function replaced by its estimated counterpart from step 1.

Recall that $p(\mathbf{X})$ has two components, the probability of each studio j choosing week $\tau^{-1}[w]$, $[p_{j\tau^{-1}[w]}(\mathbf{X})]'_{j=1,\ldots,J}$, and the probability of a market/release structure $\mathbf{R}$, denoted $p(\mathbf{R}|\mathbf{X})$. I estimate $[p_{j\tau^{-1}[w]}(\mathbf{X})]'_{j=1,\ldots,J}$ and $p(\mathbf{R}|\mathbf{X})$ separately via multinomial logit. When sample size is large enough, $p(\mathbf{X})$ can be estimated non-parametrically.

Some potential estimation difficulties are worth noting. As the dimension of firms' choice space and movie characteristics $\mathbf{X}$ increases, "curse of dimensionality" issue in estimation of $p(\mathbf{X})$ arises. For example, seasons in my model are either seven weeks or four weeks. Obtaining precise estimate of $p_{j\tau^{-1}[w]}(\mathbf{X})$ for up to six studios over all possible release weeks is demanding on sample size even when adopting a parametric approach. A similar problem applies to the estimation of market structure probability, $p((\mathbf{R}|\mathbf{X})$. Therefore, I propose a systematic and tractable way to reduce the dimension of firm choices.

First, I consolidate the week choices in a season into say two categories: weeks where underlying demand is high (H), and weeks where underlying demand is low (L). To do this, I rank all weeks in a season by their underlying demand, as measured by the average opening weekend revenue over nineteen years of data. In each season, I assign H to the highest opening weekend revenue week. I also assign H to the second highest opening weekend revenue week if its opening weekend revenue is within one standard deviation of the week with the highest opening weekend revenue. I assign L to all weeks whose average opening weekend revenues do not meet the criterion of one standard deviation from the highest average opening weekend revenue. This systematic classification reduces the dimension of the choice space from the number of weeks in a season to two. Therefore, estimating $[p_{j\tau^{-1}[w]}(\mathbf{X})]'_{j=1,\ldots,J}$ for all $\tau^{-1}[w]$ in a season reduces to estimating $[p_{jH}(\mathbf{X})]'_{j=1,\ldots,J}$, the probability of blockbuster j being released in a week with high underlying demand, conditioning on all blockbusters' observable characteristics. Robustness checks confirm that estimation results are not sensitive to this particular characterization of high and low demand weeks.

Second, as noted in the identification argument, the entry probability of every studio is a function of all movies' observed characteristics $\mathbf{X}$. Therefore, the dimension of $\mathbf{X}$ may be large when there are many firms. Furthermore, estimation of probabilities using a multinomial logit where many explanatory variables are binary may lead to "lumpy" estimates that are close to 0 or 1. To get around this issue and the curse of dimensionality, I use studio j's own X_j, and the median of competitors' X_{-j}, as regressors in the multinomial logit regression for the estimation of studio entry probabilities.[15]

Table 7 reports summary statistics on the estimated $[\hat{p}_{jH}(\mathbf{X})]'_{j=1,\ldots,J}$ and $\hat{p}((\mathbf{R}|\mathbf{X})$ across the samples. I obtain values well within 0 and 1, for both probabilities.

With $\hat{p}(\mathbf{X})$ estimated from Step 1, I now estimate the partial linear equation (4), replacing $p(X)$ by $\hat{p}(\mathbf{X})$. Note that two endogeneities are present: regarding advertising, and the choice set. For illustration, I estimate three separate models:

M1: Estimation of demand equation (1), without correcting for the endogeneities of advertising and choice set,

M2: Estimation of demand equation (1) via Two Stage Least Squares (2SLS), instrumenting for endogenous advertising with budget,

M3: Estimation of demand equation (4), via Two Stage Least Squares (2SLS), instrumenting

---

[15] The movie release game is an entry game whereby there are multiple week choices for each studio. In the more common case where the entry decision of firms is binary, such as K-Mart and Walmart deciding whether to enter a new area, or United and Delta deciding whether to offer a new route, estimation of entry probabilities is more straightforward. If data size permits, I can use more categories of week ranking, and/or use nonparametric estimators.

for endogenous advertising with budget, with a control function in equation (4) correcting for the endogenous choice set.

It is worth noting that M1 resembles Einav (2007). The difference is that the demand in Einav (2007) is modeled via a nested logit, whereby there is a nest for the inside options, and one for the outside option. The presence of a nest coupled with logistic shares introduce a term involving market shares on the right hand side of the demand equation, requiring an instrument. The author uses the number of movies in a week as the instrument. However, this instrument is no longer valid when the choice set is endogenous.

The movie characteristics X include dummy variables for three of the four genres, and "fame". Recall that "fame" indicates whether the blockbuster is a sequel, prequel, remake, and/or based off famous books, TV shows, video games, and/or historical events/characters. Much of the movie demand estimation literature has used "sequel", however, my variable "fame" appears more suitable in capturing pre-existing awareness, hype and excitement about a movie. I perform robustness checks whereby I include instead, dummy variables for sub-categories of "fame", such as sequel and remake. The key estimation results hold under these alternative specifications. In addition to movie characteristics, I also include a trend of average quarterly US household income. Its coefficient has the interpretation of the marginal utility of income, much like the negative of the price coefficient in a standard BLP model.

In the estimation of Model 3, I expand the control function $\lambda^{\xi}_{j:\mathbf{R}}[p(\mathbf{X})]$ by a linear sieve. Theoretically, any complete basis in a suitable Hilbert space of functions can be used, but in practice amongst a finite sample, the choice of basis matters. I employ a basis which economizes on parameters, and reduces to the classic Heckman correction when the movie qualities ξ_j, $j = 1, ..., J$ are independently and normally distributed. Specifically,

$$\begin{aligned}\lambda^{\xi}_{j:\mathbf{R}}[p(\mathbf{X})] =&\delta_{00} + \frac{\prod_{j=1}^{J} \phi(\Phi^{-1}(1 - p_{jH}(\mathbf{X})))}{p(\mathbf{R}|\mathbf{X})}\left[\delta_0 + \delta_1 \Phi^{-1}(1 - p_{1H}(\mathbf{X})) + ...\right.\\ &\left. + \delta_J \Phi^{-1}(1 - p_{JH}(\mathbf{X})) + ...\right], \end{aligned} \tag{5}$$

where $\Phi(\cdot)$ (resp. $\phi(\cdot)$) is the cdf (resp. pdf) of a standard Normal.

To implement the estimation of the demand equation, I truncate the polynomial series in (5) to linear terms with coefficients $(\delta_0,...,\delta_J)$, and use the probabilities estimated in Step 1 in place of $[p_{jH}(\mathbf{X})]'_{j=1,\dots,J}$ and $p(\mathbf{R}|\mathbf{X})$.[16] I then perform several statistical tests to demonstrate the performance of the basis I use.

6 Empirical Results and Counterfactuals

Table 8 presents estimation results for selected co-variates from M1, M2 and M3, and the associated adjusted R^2.[17] All coefficient estimates on key co-variates in M2 and M3 are statistically significant at 1 percent level, and all three models have adjusted R^2 close to 60 per cent.

A few things stand out. The first one is with regards to advertising: M1 underestimates the effect of advertising on box office relative to M2, and M2 in turn underestimates the effect of advertising on box office relative to M3. Recall that M1 does not account for the endogeneity of advertising, whilst M2 does. Indeed, an endogeneity test strongly rejects the null hypothesis that advertising is exogenous (see Appendix for more details). Standard instrumental variable theory says that if the regressor A_j is negatively correlated with the error term ξ_j, then not addressing this endogeneity will lead to a negative bias in the coefficient estimate. The fact that the bias on the A_j coefficient in M1 relative to M2 is negative implies that the correlation between A_j and ξ_j is negative. A Hausman test on the advertising coefficient estimates of M1 and M2 confirm that they are indeed statistically different at one per cent level of significance. On one hand, movies with high quality may choose to advertise more to signal such quality. On the other hand, they may choose to advertise less, if they are confident that critics' reviews leading up to their movie release will be positive, so not as much advertising is needed to achieve box office targets, since advertising is costly. My results above show that the latter effect dominates: movies with higher privately observed quality ξ_j spend less on advertising, and studios "compensate" for low privately observed qualities with higher advertising budgets. This finding is consistent with the hypothesis that opening week box office is highly correlated with overall success of a movie, and one way studios try to maximize opening week box office when they are aware that their movie

[16] I include all week dummies τ_t, therefore omitting the intercept term in the demand equation. PVY (2019) show that the intercept term in the demand equation is not identified, if it is firm-specific (i.e. the brand effect), when there is endogenous choice set.

[17] The estimates of week dummies (in all three models) and coefficients on the control function (in M3) are available upon request.

quality is low, is to ramp up advertising leading up to the release, before word of mouth spreads about the true quality of the movie.

Without accounting for the endogenous choice set, M2 underestimates the effect of advertising relative to M3. A Hausman test (see Appendix for more details) on the advertising coefficient estimates of M2 and M3 rejects the null hypothesis that they are statistically no different at one per cent level of significance. This finding is not surprising, given the conclusion in the previous paragraph about the negative correlation between advertising and movie quality. As discussed in the identification argument, there exists a selection on ξ_j as a result of the endogenous choice set, namely that blockbusters with higher ξ_j will choose to release in weeks with higher underlying demand. Ignoring this selection, coupled with the negative correlation between A_j and ξ_j, leads to underestimating the effect of advertising on box office.

With some algebra, one can show that the elasticity of market share with respect to own advertising, $e_{s_j,A_j} \equiv \frac{\partial s_j}{\partial A_j}\frac{A_j}{s_j}$, is equal to $\alpha(1-s_j)$. Detailed derivation is in the Appendix. I report the mean of $\hat{\alpha}(1-s_j)$ for the above estimated three models in Table 9. Indeed, the significant differences in the estimates of $\hat{\alpha}$, the effect of advertising on box office, carry over to the elasticities.

The second important observation is with regard to the fame dummy. Not accounting for the endogeneity of advertising overestimates the coefficient on fame by 19 per cent. Without correcting the advertising endogeneity (M2 versus M1), we underestimate the effect of advertising on box office, and at the same time, overestimate the effect of being famous on box office. Further taking into account the endogenous choice set increases the coefficient estimate of fame by 3 per cent. Pairwise Hausman tests on the fame coefficient estimates confirm that they are indeed statistically different from one another at one per cent level of significance. The counts of movies with fame releasing in high demand weeks versus low demand weeks are not statistically different in my sample. That is, some movies with high fame are released in high demand weeks, some are released in low demand weeks. But we know from the selection issue that the ones released in high demand weeks are likely due to having a high quality, ξ_j. If movies with fame are high quality and released in high demand weeks, their box office will be high. On the other hand, movies with fame but low quality, released in low demand weeks, will likely not to perform as well in the box office. Therefore, naive estimation without accounting for this selection may underestimate the effect of being famous on box office, as it wrongly attributes some of the mediocre performance on fame, as opposed to having a low ξ_j.

The coefficient on quarterly income is positive and significant in all three movies, confirming that going to the movie theater is a normal good. The weekly decay, as captured by the coefficient on $t - \tilde{R}$, is positive and significant, consistent with the dramatic decrease in weekly box office as time goes by after the movie's release. The negative coefficient estimates on the three genre dummies, Action/Adventure, Comedy/Family and Drama/Thriller are not to say that movies with one of these genres perform worse in the box office relative to the omitted genre, Other. The left hand side of the demand equation is $log(s_{jt}/s_{0t})$, and not s_{jt}. Therefore, the ranking of s_{jt} is not necessarily preserved when viewed relative to the outside option. In the data, it is indeed the case that the average share s_{jt} for Action/Adventure movies is larger than those for other genres, while the average $log(s_{jt}/s_{0t})$ for Action/Adventure movies are less than the other genre counterparts, implying that the share of the outside option, s_{0t}, tends to be relatively larger in weeks where Action/Adventure movies are present. Potential explanations include that Action/Adventure movies tend to dominate the summer seasons which are longer than the winter ones, therefore the outside option in any week is likely higher due to releases being less clustered. The availability of more non-blockbuster movies may also be a potential explanation. Therefore, the coefficients on the genre dummies are not direct indications of the relative impacts each genre has on box office.

Robustness

I perform various robustness checks. First, in Step 1 of the estimation of M3 with control function, I re-estimate the entry probabilities with alternative definitions of low and high demand weeks. I use both a more strict and more lenient definition of high-demand-week compared to that used in my baseline M3. My estimation results are not sensitive to these definitions. Second, I estimate all three models using alternative measures of a movie's pre-existing fame. Recall that the dummy variable "fame" used in my baseline models indicates whether the blockbuster is a sequel, prequel, remake, and/or based off famous books, TV shows, video games, and/or historical events/characters. I estimate M1 - M3 using instead, "sequel", "remake", and/or a combination. My results are robust to these alternative specifications. Finally, I estimate M1-M3 with alternative definitions of market size (moviegoer population).

The comparison of the sequel coefficient estimates in the three models are worth mentioning, I report these results in Table 10. First, the adjusted R^2 for all three models are respectively slightly lower compared to their counterparts using "fame". This confirms that the above-defined "fame" is a better measure of a movie's pre-existing fame as opposed to sequel, and that using sequel may miss out important movies such as *Titanic*, which deserves special consideration. Second, comparing the coefficient estimate of "sequel" in M1 versus M2 reveals that the effect of being sequel on a blockbuster's box office revenue is overestimated in M1 when not accounting for advertising endogeneity.[18] This suggests that some effect of advertising on box office may be wrongly attributed to the movie being a sequel, if not accounting for the endogeneity of advertising. Third, coefficient estimate of "sequel" in M2 is higher than that in M3, and again statistically different at one per cent level. In my blockbuster dataset, around one third of the blockbusters are released in a high-demand week of a season, whereas around one half of sequels are released in a high-demand week. That is, sequels are more likely to be released in a high-demand week. Therefore, not taking into account the endogenous choice set issue may lead to overestimation of the impact of being a sequel on blockbusters' box office performance.

The measure of market size, in this case, moviegoer population, is important in discrete choice demand modeling, as it determines the relative magnitudes of the market shares of inside goods relative to the outside good. I discussed my data source and how I arrived at my measure of US moviegoer population in the Data section. Previous analysis such as Einav (2007) use total US population as the market size. Such measure may overestimate the size of the market, as there are groups that will not go to the movies at all, such as infants and young children, as well as the very elderly. My measure of moviegoer population is about two thirds of total US population, which is nonetheless high, since it is unlikely that in any week, two thirds of the people in the US are potential moviegoers. The population measure I use in the main models are half of the total US moviegoer population, which is about one third of total US population, and on average twice the total number of tickets sold in Christmas week. This is a more reasonable measure of potential moviegoer population in a given week, and still a generous measure, since that population is one of the highest in Christmas week. For robustness checks I re-estimate M1-M3 using the full US moviegoer population, as well as one quarter of US moviegoer population, as alternative market size measures. Tables 11 and 12 show that indeed, key estimation results

[18]The estimates are statistically different, at 1 per cent level of significance, when performing a pair-wise Hausman Test.

are robust to these alternative measures.

6.1 Counterfactuals

This section presents several counterfactual simulations, conducted using the structural estimates from the Estimation section above. Specifically, I seek answers to the following questions: what is the cost of "non-coorperative" release compared to cooperative release where studios share their private information and maximize joint profits? What profits can studios achieve under cooperation? What would be the release weeks? Finally, inspired by the recent Walt Disney and 20th Century Fox merger, I present some analysis on the impact on release dates and profits resulting from such a merger. Key findings from these counterfactuals are presented in Tables 13.

To find the cooperative release patterns that maximize joint studio profit in a given season, I simulate profits of each studio, for all possible release choices of own and rivals'. Note that the cooperative case coincides with one where all studios share their private information, and maximize joint profits. For tractibility, I hold advertising choices unchanged from the strategic case in the data. Therefore, the profit estimates simulated can be viewed as a lower bound on profits that can be achieved through joint profit maximization. I find that on average across all seasons, joint profits increase by $25.2 millions, compared to joint profits simulated for the actual release dates chosen.[19] Furthermore, I find that when maximizing joint profits, release dates become less clustered around key weeks, on average. Specifically, going from the cooperative case to the strategic case, the number of seasons whereby there is at least one head-on-head release increases by 30 per cent.[20]

It may be more informative to look at one season, and see how the release dates in the cooperative case differs from the strategic equilibrium case. Take for example the Christmas of 2013. Fox, Paramount, and Universal released *The Secret Life of Walter Mitty, The Wolf of Wall Street* and *47 Ronin* all in week 52, the week of Christmas, while Warner released *The*

[19]Note that I can calculate a studio's actual profit by summing its weekly box office revenue, then taking away advertising expenditure. However, for the purpose of comparing with the simulated profits of joint profit maximization, it is more consistent to use the profits simulated for the actual release choices (sum of product of weekly market size and market shares, the latter computed using observable characteristics and estimated structural parameters, minus advertising expenditure).

[20]By head-on-head release, I mean that at least two new releases in the same week.

Hobbit in week 50. In contrast, the cooperative outcome has two of the blockbusters releasing in week 49, one in week 51 and one in week 52. Joint profits increase to $496 million, from $493 million in the strategic case.

There is indeed evidence that in the case of strategic release, studios choose to release in the high-demand weeks (weeks containing major holidays) more frequently, compared to the cooperative release that maximizes joint profits. The above season is an example of this. Take another season, Christmas season of 2015, Fox, Disney and Warner released *The Revenant, Star Wars Ep. VII: The Force Awakens* and *Point Break* respectively in weeks 52, 51 and 52. In the cooperative outcome however, Fox would release in week 49, Disney in week 52 and Warner in week 51. This leads to a joint profit of 629 million, compared to a joint profit of $618 million in the strategic case. Furthermore, consumer surplus increases from $457 million in the strategic release case to $463 million in the cooperative release case. More variety is better for consumers, and in the context of movies, this translates to having more spread-out release dates. In addition, a more spread-out release means in a given week of the season, there is likely to be a greater number of movies.

Finally, in the counterfactual where Fox and Disney merge, I compute an upper bound on their joint profits after merger. I find that jointly, the merged firms can achieve on average a maximum of $217 million profits. This is $20.6 million higher than their joint profits under the cooperative scenario, and $31.6 million higher than their joint profits under the non-coorperative scenario.

To better understand the impact of the merger, I look at the impact in a specific season. Again, take the example of Christmas season of 2015, Fox, Disney and Warner released *The Revenant, Star Wars Ep. VII: The Force Awakens* and *Point Break* respectively in weeks 52, 51 and 52. My analysis suggest that after merger, the Fox-Disney profits increase to $642 million, compared to $631 million pre-merger. In contrast, Warner's profit decreases by $0.7 million. My model predicts that Fox-Disney would release their two blockbusters in weeks 49 and 52 respectively, as opposed to 52 and 51 in the strategic case. Indeed, having the two blockbusters more apart means they compete less with one another. Since more movies are released early in the season, it also means that there are more blockbuster choices per week of the season.

The merger has some important implications that differ from a typical horizontal merger. The merger took 6 months to obtain the approval of the Justice Department, much less than the average time taken for a horizontal merger, especially one of this scale. According to `variety.com`, the Justice Department's main concern regarding the merger was that Disney's purchase of the Regional Sports Networks (RSNs) would "harm competition for sports programming in those local markets". Thus, Disney's agreement to divest the twenty-two RSNs paved ways for the merger's approval. In addition, unlike the typical horizontal merger, a hike in movie ticket is not a concern for the Justice Department, given the motion pictures industry's uniform movie ticket pricing. Indeed, I find that in the Christmas 2015 season, there is a $5.4 million increase in consumer surplus post-merger, due to the fact that release dates became more spread out in the season. There is minimal change to the consumer surplus compared to the cooperative release case. A full summary of analyses for the Christmas 2015 season is in Table 14.

However, the merged Disney-Fox will likely have higher market power when it comes to negotiating profit share and screening slots with theaters. Disney is already known for its aggressive negotiations with theaters for its big blockbusters such as Star Wars. The merger will see its market share increase to as much as 30 per cent of the industry, or over 35 per cent amongst of the "Big Six", which have now become the "Big Five". I cannot analyze this in light of my model estimates, as the latter are based on estimation of 6-studio oligopolist model, where each studio is an equal player. However, it may be expected that Disney-Fox will be able to secure more favorable release weeks and screening slots for their future movies, as well as negotiate a higher share of revenue from screenings.

7 Conclusion

This paper performs an analysis of strategic release dates and advertising of blockbuster movies. In each week of a release season, consumers choose to watch a blockbuster movie or some outside option, in order to maximize utility. Studios, knowing how consumers behave, need to make optimal release and advertising decisions to maximize expected profits, amid rival blockbusters' unobserved qualities. Incorporating studios' strategic play introduces two endogeneity issues: namely the endogenous nature of advertising, and the endogenous choice set in each week as a result of studios' strategic release timing.

My main findings include significant underestimation of the effect of advertising on box office when ignoring one or both endogeneities, due to the negative correlation between unobserved own movie quality and advertising, and/or selection on quality. On the other hand, the impact of being a sequel on box office is overestimated, when ignoring endogenous choice sets, as sequels are more likely to be released in high-demand weeks.

The methods used in this paper can be applied to many other industries where strategic release and advertising (or pricing) are important on the supply side. Indeed, endogenous choice sets are present in many oligopoly industries, such as airlines choosing which routes to cover, hotels picking which areas to locate, food and beverage brands choosing where and which stores to sell, and supermarkets deciding whether to open a store in a new area. My estimation framework can be readily adapted to these applications. Future research may look at extending the demand model to one of a nested discrete choice model,[21] separating advertising decisions into various media categories,[22] and expanding the modeling framework to multiproduct firms.

A Appendices

A.1 Supply Model for J Studios and W Week Choices

There are now W^J possibilities of release patterns:

$$(1,1,...,1),(1,1,...,2),...,...,(W,W,...,W)$$

Case $(R_1, R_2, ..., R_J) = (1,1,...,1)$:

This is the case whereby all J studios choose to release in week 1. The profit of studio 1 in this case is:

$$\pi_{1,(1,1,...,1)} = M_1 s_{1,1} + M_2 s_{1,2} + M_3 s_{1,3} ... + M_8 s_{1,8} - A_1 = \sum_{t=1}^{8} M_t s_{1,t} - A_1$$

Note that in the above expression, it ends in M_8 because we still assume that each movie shows for 8 weeks. This profit expression looks identical to those for the 2-studio 2-week case detailed in the model section, with the shares computed appropriately. It is straightforward to show that

[21] Although the availability of a good instrument is a challenge.

[22] The data source of advertising I used, Ad$ummary/Ad$pender, of Kantar Media, only started publishing advertising spending by media categories since 2009 (my box office data begin in 1997).

for any possible release $\tilde{\boldsymbol{R}} = (\tilde{R}_1, ..., \tilde{R}_J)$,

$$\pi_{j,\tilde{\mathbf{R}}} = \sum_{t=\tilde{R}_j}^{\tilde{R}_j+7} M_t s_{j,t} - A_j$$

These are now shares from monopoly (only one studio j out of all J studios are present in a week) up to J-studio-oligopoly (all J studios' blockbusters are present that week) shares:

$$\begin{aligned} s_{jt}^m &= \frac{\exp^{X_j'\beta - \lambda(t-\tilde{R}_j)+\alpha \log(A_j)+\xi_j+\eta_{jt}}}{\exp^{-\tau_t} + \exp^{X_j'\beta - \lambda(t-\tilde{R}_j)+\alpha \log(A_j)+\xi_j+\eta_{jt}}} \\ &= s_{jt}^m(\tilde{R}_j, A_j, \xi_j) \\ s_{jt}^{oligopoly} &= \frac{\exp^{\mu_j+\tau_t-\lambda(t-\tilde{R}_j)+\alpha \log(A_j)+\xi_j}}{\exp^{-\tau_t} + \sum_{j'\in C_t} \exp^{X_{j'}'\beta - \lambda(t-\tilde{R}_{j'})+\alpha \log(A_{j'})+\xi_{j'}+\eta_{j't}}} \\ &= s_{jt}^{oligopoly}(\tilde{R}_{j'}, A_{j'}, \xi_{j'} : j' \in C_t) \end{aligned}$$

where C_t denotes the set of available blockbusters in week t. Note that for any set of release decisions $\tilde{R}_1, ..., \tilde{R}_J$, we can directly infer the choice set (may be empty) in each week of the season, C_t, $t = 1, ..., W$.

2nd stage: Advertising Decision

Each of the J studios chooses their advertising level to maximize their total 8-week profit, taking the other $J-1$ studios' advertising levels A_{-j} as given. Studios play a Bertrand-style game on advertising, choosing their optimal advertising by best-responding to the advertising levels of its rivals under incomplete information regarding ξ_{-j} and η_{jt}. This leads to J first order conditions:

$$\frac{\partial \mathbb{E}[\pi_{j,\mathbf{R}}|\xi_j, \mathbf{R}]}{\partial A_j} = 0, \ j = 1, ..., J$$

Solving for $(A_1, ..., A_J)$ simultaneously from the above J first order conditions yields the optimal advertising strategies:

$$A_j = A_j^*(\mathbf{X}, \xi_j), \ j = 1, ..., J$$

and the corresponding maximized profits:

$$\pi_{j,\mathbf{R}}^*(\mathbf{A}^*) = \pi_{j,\mathbf{R}}^*(\mathbf{X}, \boldsymbol{\xi}), \ j = 1, ..., J$$

1st stage: Release Date Decision

I rank weeks 1 through W, $w = 1, ..., W$, and define $\tau^{-1}[w]$ the same way as in the 2-studio 2-week case: let $\tau^{-1}[W]$ as period with the lowest underlying demand τ_t, $\tau^{-1}[W-1]$ as period with the second lowest underlying demand,..., and $\tau^{-1}[1]$ as period with the highest underlying demand.

I assume monotone Bayesian Nash equilibrium strategies. Given that there may be more than two weeks in a season now, there will be more than one equilibrium threshold. Specifically, there will be $W-1$ equilibrium thresholds $\xi_j^*[1]$,...,$\xi_j^*[W-1]$ that make studio $j = 1, ..., J$ indifferent between releasing in week $\tau^{-1}[1]$ and week $\tau^{-1}[2]$, ..., and between releasing in week $\tau^{-1}[W-1]$ and week $\tau^{-1}[W]$. That is, for $j = 1, ..., J$,

$$R_j = r(\xi_j) = \begin{cases} \tau^{-1}[1] & \text{if } \xi_j \geq \xi_j^*[1] \\ \tau^{-1}[2] & \text{if } \xi_j^*[2] \leq \xi_j < \xi_j^*[1] \\ \vdots & \\ \tau^{-1}[W] & \text{if } \xi_j < \xi_j^*[W-1], \end{cases}$$

where $\xi_j^*[t]$, $j = 1, ..., J, t = 1, ..., W-1$ solve simultaneously:

$$\begin{aligned} &\mathbb{E}\left[\pi^*_{1,(\tau^{-1}[1],R_2,...,R_J)}(\xi_1^*[1], \xi_2, ..., \xi_J)|\xi_1^*[1]\right] \\ = &\mathbb{E}\left[\pi^*_{1,(\tau^{-1}[2],R_2,...,R_J)}(\xi_1^*[1], \xi_2, ..., \xi_J)|\xi_1^*[1]\right], \end{aligned}$$

$$\begin{aligned} &\mathbb{E}\left[\pi^*_{2,(R_1,\tau^{-1}[1],R_3,...,R_J)}(\xi_1, \xi_2^*[1], \xi_3, ..., \xi_J)|\xi_2^*[1]\right] \\ = &\mathbb{E}\left[\pi^*_{2,(R_1,\tau^{-1}[2],R_3,...,R_J)}(\xi_1, \xi_2^*[1], \xi_3, ..., \xi_J)|\xi_2^*[1]\right], \\ &\vdots \\ &\mathbb{E}\left[\pi^*_{J,(R_1,...,\tau^{-1}[1])}(\xi_1, ..., \xi_J^*[1])|\xi_J^*[1]\right] \\ = &\mathbb{E}\left[\pi^*_{J,(R_1,...,\tau^{-1}[2])}(\xi_1, ..., \xi_J^*[1])|\xi_J^*[1]\right]; \\ &\vdots \\ &\vdots \\ &\mathbb{E}\left[\pi^*_{1,(\tau^{-1}[W-1],R_2,...,R_J)}(\xi_1^*[W-1], \xi_2, ..., \xi_J)|\xi_1^*[W-1]\right] \\ = &\mathbb{E}\left[\pi^*_{1,(\tau^{-1}[W],R_2,...,R_J)}(\xi_1^*[W-1], \xi_2, ..., \xi_J)|\xi_1^*[W-1]\right], \end{aligned}$$

$$
\begin{aligned}
&\mathbb{E}\left[\pi^*_{2,(R_1,\tau^{-1}[W-1],R_3,...,R_J)}(\xi_1,\xi_2^*[W-1],\xi_3,...,\xi_J)|\xi_2^*[W-1]\right]\\
&=\mathbb{E}\left[\pi^*_{2,(R_1,\tau^{-1}[W],R_3,...,R_J)}(\xi_1,\xi_2^*[W-1],\xi_3,...,\xi_J)|\xi_2^*[W-1]\right],\\
&\qquad\vdots\\
&\mathbb{E}\left[\pi^*_{J,(R_1,...,\tau^{-1}[W-1])}(\xi_1,...,\xi_J^*[W-1])|\xi_J^*[W-1]\right]\\
&=\mathbb{E}\left[\pi^*_{J,(R_1,...,\tau^{-1}[W])}(\xi_1,...,\xi_J^*[W-1])|\xi_J^*[W-1]\right].
\end{aligned}
$$

Similar to the 2-firm 2-week case, we can express the expectations in the above indifference conditions as intergrals, for example

$$
\mathbb{E}\left[\pi^*_{1,(\tau^{-1}[1],R_2,...,R_J)}(\xi_1^*[1],\xi_2,...,\xi_J)|\xi_1^*[1]\right]=\mathbb{E}\left[\pi^*_{1,(\tau^{-1}[2],R_2,...,R_J)}(\xi_1^*[1],\xi_2,...,\xi_J)|\xi_1^*[1]\right],
$$

as:

$$
\begin{aligned}
&Pr\left(\xi_2\le\xi_2^*[W-1],...,\xi_J\le\xi_J^*[W-1]|\xi_1^*[1]\right)\int_{-\infty}^{\infty}...\int_{-\infty}^{\infty}\int_{-\infty}^{\xi_2^*[W-1]}...\int_{-\infty}^{\xi_J^*[W-1]}\\
&\quad\pi^*_{1,(\tau^{-1}[1],\tau^{-1}[W],...,\tau^{-1}[W])}(\xi_1^*[1],\xi_2,...,\xi_J)dF_{\xi_2,...,\xi_J|\xi_1^*[1]}(\xi_2,...,\xi_J|\xi_1^*[1])dF_{\eta_1}...dF_{\eta_J}\\
&+Pr\left(\xi_2^*[W-1]<\xi_2\le\xi_2^*[W-2],...,\xi_J\le\xi_J^*[W-1]|\xi_1^*[1]\right)\int_{-\infty}^{\infty}...\int_{-\infty}^{\infty}\int_{\xi_2^*[W-2]}^{\xi_2^*[W-1]}...\int_{-\infty}^{\xi_J^*[W-1]}\\
&\quad\pi^*_{1,(\tau^{-1}[1],\tau^{-1}[W-1],\tau^{-1}[W],...,\tau^{-1}[W])}(\xi_1^*[1],\xi_2,...,\xi_J)dF_{\xi_2|\xi_1^*[1]}(\xi_2,...,\xi_J|\xi_1^*[1])dF_{\eta_1}...dF_{\eta_J}\\
&\qquad\vdots\\
&+Pr\left(\xi_2\ge\xi_2^*[1],...,\xi_J\ge\xi_J^*[1]|\xi_1^*[1]\right)\int_{-\infty}^{\infty}...\int_{-\infty}^{\infty}\int_{\xi_2^*[1]}^{\infty}...\int_{\xi_J^*[1]}^{\infty}\\
&\quad\pi^*_{1,(\tau^{-1}[1],...,\tau^{-1}[1])}(\xi_1^*[1],\xi_2,...,\xi_J)dF_{\xi_2,...,\xi_J|\xi_1^*[1]}(\xi_2,...,\xi_J|\xi_1^*[1])dF_{\eta_1}...dF_{\eta_J}\\
=\ &Pr\left(\xi_2\le\xi_2^*[W-1],...,\xi_J\le\xi_J^*[W-1]|\xi_1^*[1]\right)\int_{-\infty}^{\infty}...\int_{-\infty}^{\infty}\int_{-\infty}^{\xi_2^*[W-1]}...\int_{-\infty}^{\xi_J^*[W-1]}\\
&\quad\pi^*_{1,(\tau^{-1}[2],\tau^{-1}[W],...,\tau^{-1}[W])}(\xi_1^*[1],\xi_2,...,\xi_J)dF_{\xi_2,...,\xi_J|\xi_1^*[1]}(\xi_2,...,\xi_J|\xi_1^*[1])dF_{\eta_1}...dF_{\eta_J}\\
&+Pr\left(\xi_2^*[W-1]<\xi_2\le\xi_2^*[W-2],...,\xi_J\le\xi_J^*[W-1]|\xi_1^*[1]\right)\int_{-\infty}^{\infty}...\int_{-\infty}^{\infty}\int_{\xi_2^*[W-2]}^{\xi_2^*[W-1]}...\int_{-\infty}^{\xi_J^*[W-1]}\\
&\quad\pi^*_{1,(\tau^{-1}[2],\tau^{-1}[W-1],\tau^{-1}[W],...,\tau^{-1}[W])}(\xi_1^*[1],\xi_2,...,\xi_J)dF_{\xi_2|\xi_1^*[1]}(\xi_2,...,\xi_J|\xi_1^*[1])dF_{\eta_1}...dF_{\eta_J}\\
&\qquad\vdots\\
&+Pr\left(\xi_2\ge\xi_2^*[1],...,\xi_J\ge\xi_J^*[1]|\xi_1^*[1]\right)\int_{-\infty}^{\infty}...\int_{-\infty}^{\infty}\int_{\xi_2^*[1]}^{\infty}...\int_{\xi_J^*[1]}^{\infty}\\
&\quad\pi^*_{1,(\tau^{-1}[2],\tau^{-1}[1],...,\tau^{-1}[1])}(\xi_1^*[1],\xi_2,...,\xi_J)dF_{\xi_2,...,\xi_J|\xi_1^*[1]}(\xi_2,...,\xi_J|\xi_1^*[1])dF_{\eta_1}...dF_{\eta_J}.
\end{aligned}
$$

A.2 Elasticity Derivations

A.2.1 Own elasticity of market share w.r.t. advertising

$$
\begin{aligned}
\frac{\partial s_j}{\partial A_j} &= \frac{\partial s_j}{\partial \log(A_j)} \frac{\partial \log(A_j)}{\partial A_j} \\
&= \frac{\partial s_j}{\partial \log(A_j)} \frac{1}{A_j} \\
&= \frac{\partial \dfrac{\exp^{X_j'\beta - \lambda(t-\tilde{R}_j) + \alpha \log(A_j) + \xi_j + \eta_{jt}}}{\exp^{-\tau_t} + \sum_{j' \in C_t} \exp^{X_j'\beta - \lambda(t-\tilde{R}_j) + \alpha \log(A_j) + \xi_j + \eta_{jt}}}}{\partial \log(A_j)} \frac{1}{A_j} \\
&= \alpha(s_j - s_j^2) \frac{1}{A_j}
\end{aligned}
$$

Therefore, $e_{s_j,A_j} \equiv \frac{\partial s_j}{\partial A_j} \frac{A_j}{s_j} = \alpha(1 - s_j)$

A.2.2 Cross elasticity of market share w.r.t. advertising

$$
\begin{aligned}
\frac{\partial s_j}{\partial A_k} &= \frac{\partial s_j}{\partial \log(A_k)} \frac{\partial \log(A_k)}{\partial A_k} \\
&= \frac{\partial s_j}{\partial \log(A_k)} \frac{1}{A_k} \\
&= \frac{\partial \dfrac{\exp^{X_j'\beta - \lambda(t-\tilde{R}_j) + \alpha \log(A_j) + \xi_j + \eta_{jt}}}{\exp^{-\tau_t} + \sum_{j' \in C_t} \exp^{X_j'\beta - \lambda(t-\tilde{R}_j) + \alpha \log(A_j) + \xi_j + \eta_{jt}}}}{\partial \log(A_k)} \frac{1}{A_k} \\
&= -\alpha s_j s_k \frac{1}{A_k}
\end{aligned}
$$

Therefore, $e_{s_j,A_k} \equiv \frac{\partial s_j}{\partial A_k} \frac{A_k}{s_j} = -\alpha s_k$

A.2.3 Hypothesis Tests

Test for endogeneity of advertising

Regress A_j on exogenous regressors and instruments, obtain residuals. Then, include residuals as an extra regressor in original regression and perform a t-test on its coefficient. The p-value from this test is 0.000. Again, I conclude that A_j is endogenous.

F test on M3

Here I test the joint significance of all the coefficients of terms in the control function: M2 is the restricted model, and M3 is the unrestricted model. There are $q = 7$ restrictions (λ_0 - λ_6 in control function) in the null. $F = [(SSE_r - SSE_u r)/q]/[SSR_u r/(n-k-1)]$ where n is the number of observations and k is the number of regressors (not including intercept but there is no intercept terms anyway) in unrestricted model. I obtain $F = 10.3590$, with a critical value of 2.0129. Therefore, I reject the null that the λs are jointly insignificant since the F statistic is much higher than the critical value.

Hausman test

I also perform various Hausman tests to see if key coefficients in M1 versus M2, and M2 versus M3 are significantly different. I focus on the coefficients on A_j, $t - \tilde{R}$, quarterly income, genres, and fame.

Hausman test statistic is $\frac{N(\hat{\beta}_1 - \hat{\beta}_0)^2}{V(\hat{\beta}_1) - V(\hat{\beta}_0)}$, and distributed χ^2_1 where $\hat{\beta}_0$ is the estimate under the null hypothesis of no misspecification. That is, $\hat{\beta}_0$ are the estimated coefficients from M1 (resp. M2) for the comparison between M1 and M2 (resp. M2 and M3). The 5% critical value for χ^2_1 is 3.8415. Given the reported test statistics below, the null that the 2 sets of βs are different for each of the regressors, are all strongly rejected.

	$\log(A_j)$	t-$\tilde{R}$	Qtr Inc	Genre AA	Genre CF	Genre DT	Fame
Test Stat.	14996.82	1935.93	17162.42	754661.40	646326.57	1089268.47	141373.07

Table 1: Hausman Test Statistics, M1 versus M2 Comparison

	$\log(A_j)$	t-$\tilde{R}$	Qtr Inc	Genre AA	Genre CF	Genre DT	Fame
Test Stat.	81743.53	73216.24	1938.45	39693.70	56747.82	85226.60	68695.11

Table 2: Hausman Test Statistics, M2 versus M3 Comparison

Tables

(1995-2018)	Studio	Total Box Office	Share	Cumulative Share
1	Walt Disney	$35,541,085,799	16.20%	16.25%
2	Warner Bros	$33,232,824,065	15.32%	31.52%
3	Sony	$26,624,081,732	12.19%	43.71%
4	Universal	$25,232,636,854	11.64%	55.35%
5	Fox	$25,030,269,234	11.42%	66.77%
6	Paramount	$23,304,054,064	10.61%	77.38%
7	Lionsgate	$8,633,513,515	3.92%	81.30%
8	New Line	$6,193,114,702	2.79%	84.09%
9	Dreamworks	$4,278,649,271	1.93%	86.02%
10	Miramax	$3,840,594,867	1.73%	87.75%

Source: The Numbers.[23]

Table 3: Top 10 studios by grossings, 1995-2018

Ranking	Genre	Total Box Office	Share
1	Adventure	$59,427,172,759	26.93%
2	Action	$44,224,524,060	20.04%
3	Drama	$35,637,540,461	16.15%
4	Comedy	$33,070,791,998	14.99%
5	Thriller/Suspense	$18,254,310,421	8.27%
6	Horror	$10,866,490,854	4.93%
7	Romantic Comedy	$9,795,646,781	4.44%
8	Musical	$4,071,425,542	1.85%

Source: The Numbers.

Table 4: Top 8 genres by grossings, 1995-2018

[23]Dreamworks was in partnerships with Paramount, Walt Disney and Universal respectively from 2006-2008, 2009-2016 and 2016-present.

	Frequency of release
Mon	0.2 %
Tue	1.0%
Wed	9.3%
Thu	2.5%
Fri	86.6%
Sat	0.0%
Sun	0.0 %

Table 5: Proportions of releases falling on Monday - Sunday, all movies by the Big Six studios, 1997-2015

	Mean	Median	St. Dev.
Budget (excl. advertising)	$ 116,697,861	$ 99,490,862	$ 58,125,648
Domestic revenue	$ 145,727,828	$ 117,262,351	$ 116,817,559
Opening w/e rev.	$ 39,392,087	$ 29,692,605	$ 33,579,512
Advertising	$ 30,141,888	$ 30,138,805	$ 8,385,298
Advertising/Budget	0.300	0.270	0.129

Source: The Numbers, Bureau of Economic Analysis.

Table 6: Summary statistics for blockbuster movies, deflated USD, 1997-2015

	Mean	Median	St. Dev.
$[\hat{p}_{jH}(X)]'_{j=1,...,J}$	0.3689	0.3523	0.2177
$\hat{p}((R_1,...,R_J\|X)$	0.2490	0.2525	0.1420

Table 7: Summary Statistics for Estimated Probabilities of Entry and Market Structure

	M1	M2	M3
$\log(A)$	1.3548***	1.4708***	1.5509***
t-$\tilde{R}$	0.5410***	0.5425***	0.5303***
Qtr income	0.1340***	0.1427***	0.1389***
Genre AA	0.1722+	-0.6351***	-0.7328***
Genre CF	0.4270**	-0.4255***	-0.5514***
Genre DT	0.2524+	-0.5857***	-0.7785***
Fame	0.4490***	0.3766***	0.3874***
Adj Rsq	0.6000	0.5921	0.6017

* ** *** significant at 10%, 5% and 1% respectively. + not significant at 10+%.

Table 8: Estimation Results for Key Co-Variates and Model Adjusted R^2

	M1	M2	M3
e_{s_j,A_j}	1.3191	1.4320	1.5101

Table 9: Estimated Elasticity of Market Share w.r.t. Own Advertising, mean value estimates

	M1	M2	M3
$\log(A)$	1.3942***	1.4770***	1.5700***
t-$\tilde{R}$	0.5507***	0.5484***	0.5360***
Qtr income	0.1382***	0.1435***	0.1395***
Genre AA	0.0142+	-0.7443***	-0.8242***
Genre CF	0.2351+	-0.5284***	-0.6373***
Genre DT	0.1236+	-0.6218***	-0.8083***
Sequel	0.4747***	0.4651***	0.4430***
Adj Rsq	0.5983	0.5883	0.5968

* ** *** significant at 10%, 5% and 1% respectively. + not significant at 10+%.

Table 10: Estimation Results for Key Co-Variates and Model Adjusted R^2, Robustess check using "Sequel" instead of "Fame"

	M1	M2	M3
$\log(A)$	1.3506***	1.4724***	1.5524***
t-$\tilde{R}$	0.5389***	0.5404***	0.5283***
Qtr income	0.1310***	0.1400***	0.1362***
Genre AA	0.1676+	-0.6402***	-0.7355***
Genre CF	0.4197**	-0.4337***	-0.5568***
Genre DT	0.2466+	-0.5915***	-0.7828***
Fame	0.4472***	0.3744***	0.3842***
Adj Rsq	0.5931	0.5854	0.5950

* ** *** significant at 10%, 5% and 1% respectively. + not significant at 10+%.

Table 11: Estimation Results for Key Co-Variates and Model Adjusted R^2, Robustess check using total moviegoer population reported by MPAA (twice the population used in main model)

	M1	M2	M3
$\log(A)$	1.3754***	1.4764***	1.5589***
t-$\tilde{R}$	0.5477***	0.5490***	0.5365***
Qtr income	0.1454***	0.1533***	0.1495***
Genre AA	0.1946+	-0.6169***	-0.7239***
Genre CF	0.4571**	-0.3989**	-0.5355***
Genre DT	0.2732+	-0.5707***	-0.7678***
Fame	0.4503***	0.3785***	0.3937***
Adj Rsq	0.6147	0.6065	0.6160

* ** *** significant at 10%, 5% and 1% respectively. + not significant at 10+%.

Table 12: Estimation Results for Key Co-Variates and Model Adjusted R^2, Robustess check using 1/4 total moviegoer population reported by MPAA (half the population used in main model)

	M1	M2	M3
Joint profits under equilibrium ($, mil)	459.20	489.68	436.30
Joint profits under cooperative case ($, mil)	482.01	516.77	461.54
Joint profits under Disney-Fox Merger ($, mil)	462.86	481.11	443.32
Increase in joint profits SP vs equilib. ($, mil)	23.12	27.45	25.24
Max. increase in Disney-Fox profit vs cooperative ($, mil)	20.31	23.26	20.64
Max. increase in Disney-Fox profit vs equilib. ($, mil)	29.66	31.95	31.60

Joint profits refer to the sum of profits of all studios present in the season.

Table 13: Simulated Profits in Counterfactuals, mean value estimates

	Fox	Disney	Warner
Profit, strategic equilibrium ($, mil)	240.75	390.00	-12.74
Profit, cooperative release ($, mil)	203.57	439.43	-15.03
Profit, Fox-Disney merger ($, mil)	204.08	438.41	-13.43
Release week, strategic equilibrium	52	51	52
Release week, cooperative release	49	52	51
Release week, Fox-Disney merger	49	49	52

Recall that I only look at domestic box office, hence studios' overall profits would be higher.
Warner's Christmas 2015 blockbuster *Point Break*, was nonetheless a flop at the box office.

Table 14: Christmas 2015 Example

Figures

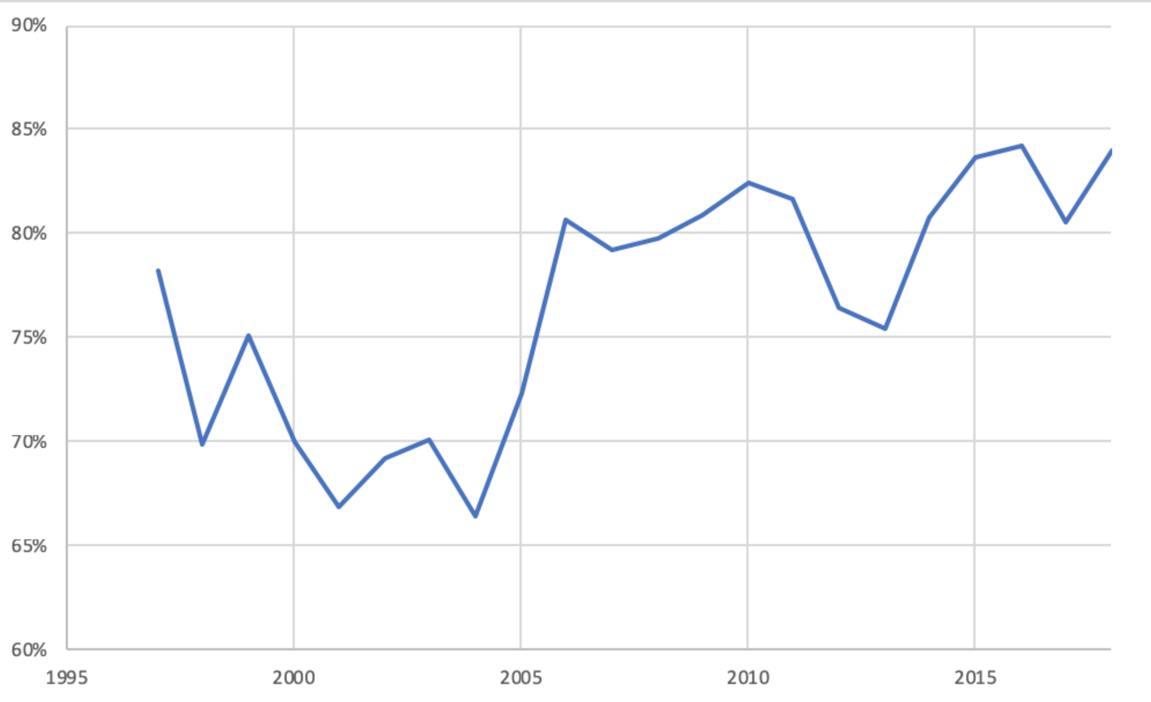


Figure 1: Annual sum of market share of the Big Six studios, 1997-2015

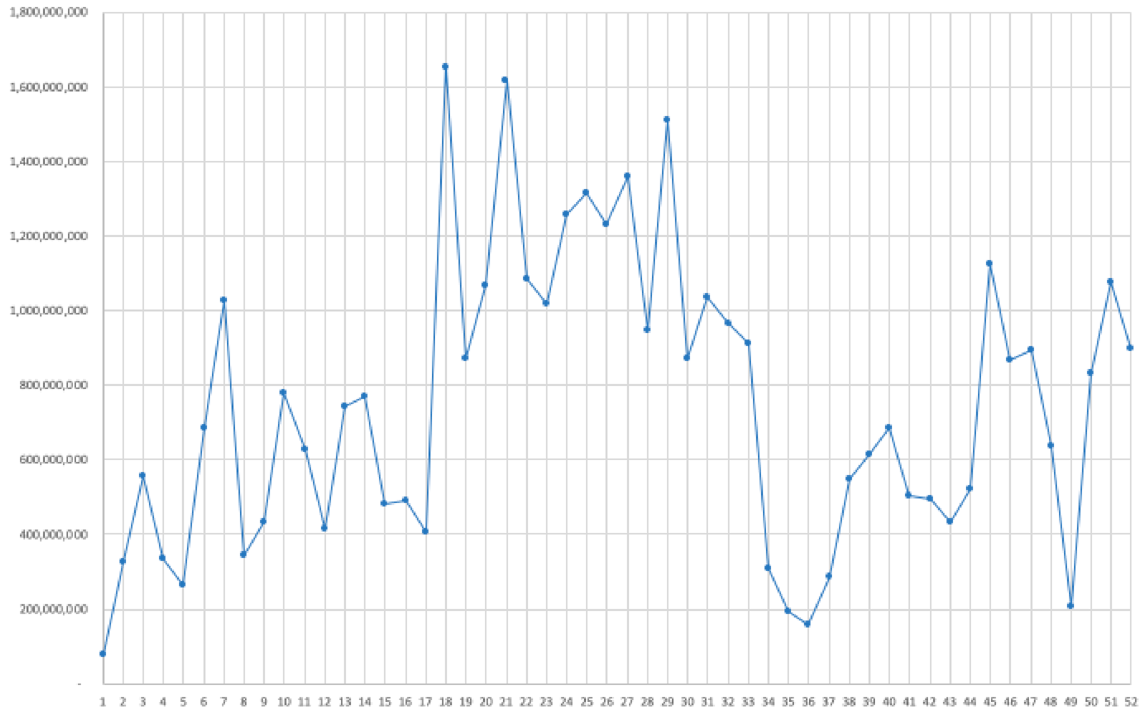


Figure 2: Inflation-adjusted opening weekend box office revenue of all Big Six studio movies, 1997-2015

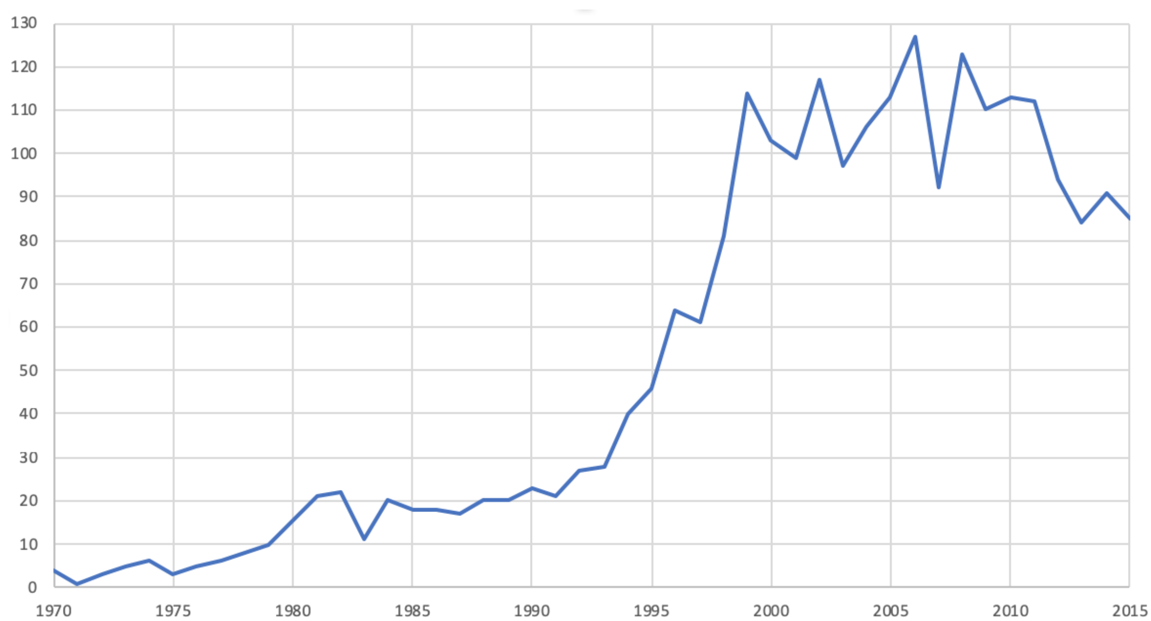


Figure 3: Annual number of movies released by the Big Six studios, 1970-2015

References

[1] Ackerberg, D. (2001): "Empirically Distinguishing Informative and Prestige Effects of Advertising", *RAND Journal of Economics*, 32, 316-333.

[2] Ackerberg, D. (2003): "Advertising, Learning and Consumer Choice in Experience Good Markets: An Empirical Examination", *International Economic Review*, 44, 1007-1040.

[3] Bagwell, K. (2007): "The Economic Analysis of Advertising", Chapter 28 in *Handbook of Industrial Organization*, Vol. 3, 1701-1844.

[4] Athey, S. (2007): "Single Crossing Properties and the Existence of Pure Strategy Equilibria in Games of Incomplete Information", *Econometrica*, Vol. 69, 861-889.

[5] Belleflamme, P. and D. Paolini (2018): "Strategic attractiveness and release decisions for cultural goods," *Journal of Economics and Management Strategy*, In Press.

[6] Berry, S., (1992): "Estimation of a Model of Entry in the Airline industry," *Econometrica*, 60, 889-917.

[7] Berry, S., (1994): "Estimating Discrete-Choice Models of Product Differentiation," *Rand Journal of Economics*, 25, 242-262.

[8] Berry, S., J. Levinsohn and A. Pakes, (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63, 841-890.

[9] Bresnahan, T. and P. Reiss, (1990): "Entry in Monopoly Markets," *Review of Economic Studies*, 57, 531-553.

[10] Bresnahan, T. and P. Reiss, (1991): "Entry and Competition in Concentrated Markets," *Journal of Political Economy*, 99, 977-1009.

[11] Cabral, L. and G. Natividad, (2016): "Box-Office Demand: the Importance of Being #1," *Journal of Industrial Economics*, 64, 277-294.

[12] Chung, H. S. (2015): "A Note on Uniform Pricing in the Motion-Picture Industry," *Hitotsubashi Journal of Economics*, 56, 231-242.

[13] Ciliberto, F., C. Murry and E. Tamer (2018): "Market Structure and Competition in Airline Markets," working paper.

[14] Compiani, G. (2019): "Market Counterfactuals and the Specification of Multi-Product Demand: A Nonparametric Approach," working paper.

[15] Courty, P (2011): "Unpriced quality," *Economics Letters*, 111, 13–15.

[16] Das, M., W. Newey and F. Vella (2003): "Nonparametric Estimation of Sample Selection Models," *Review of Economic Studies*, 70, 33-58.

[17] Daughety, A. and J. Reinganum (2008): "Imperfect Competition and Quality Signaling," *RAND Journal of Economics*, 39, 163-183.

[18] Davis, P (2005): "The Effect of Local Competition on Admission Prices in the U.S. Motion Picture Exhibition Markets," *Journal of Law and Economics*, 48, 677-707.

[19] Davis, P (2006a): "Measuring the Business Stealing Cannibalization and Market Expansion Effects of Entry in the U.S. Motion Picture Exhibition Market," *Journal of Industrial Economics*, 54, 964–982.

[20] Davis, P (2006b): "Spatial Competition in Retail Markets: Movie Theaters," *RAND Journal of Economics*, 37, 964–982.

[21] De Vany, A and W. D. Walls (1996): "Bose-Einstein Dynamics and Adaptive Contracting in the Motion Picture Industry," *The Economic Journal*, 106, 1493-1514.

[22] De Vany, A and W. D. Walls (1999): "Uncertainty in the Movie Industry: Does Star Power Reduce the Terror of the Box Office?" *Journal of Cultural Economics*, 23, 285-318.

[23] Dubé, J., G. Hitsch, and P. Manchanda (2007): "An Empirical Model of Advertising Dynamics," *Quantitative Marketing and Economics*, 3, 107-144.

[24] Einav, L. (2007): "Seasonality in the U.S. Motion Picture Industry," *Rand Journal of Economics*, 38, 127-145.

[25] Einav, L. (2010): "Not All Rivals Look Alike: Estimating an Equilibrium Model of the Release Date Timing Game," *Economic Inquiry*, 48, 369-390.

[26] Elberse, A. and J. Eliashberg (2003): "Demand and Supply Dynamics for Sequentially Released Products in International Markets: The Case of Motion Pictures," *Marketing Science*, 22, 329–354.

[27] Elberse, A. (2007): "The Power of Stars: Do Star Actors Drive the Success of Movies?," *Journal of Marketing*, 71, 102-120.

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

[29] Eliashberg, A. and S. M. Shugan (1997): "Film Critics: Influencers or Predictors?" *Journal of Marketing*, 61, 68–78.

[30] Eliashberg, J., J. Jonker, M. S. Sawhney and B. Wierenga (2000), "MOVIEMOD: An Implementable Decision-Support System for Prerelease Market Evaluation of Motion Pictures," *Marketing Science*, 19, 226–243.

[31] Eliashberg, J. (2005), "The Films Exhibition Business: Critical Issues, Practice, and Research," C. C. Moul, ed. *A Concise Handbook of Movie Industry Economics*, Cambridge University Press, 138-162.

[32] Eliashberg, J., A. Elberse and M.A.A.M. Leenders (2006), "The Motion Picture Industry: Critical Issues in Practice, Current Research, and New Research Directions," *Marketing Science*, 25, 638-661.

[33] Eliashberg, J., C. Weinberg, S. K. Hui (2008), Decision Models for the Movie Industry, Chapter 13 of *Handbook of Marketing Decision Models.*

[34] Eliashberg, J., S. K. Hui, and Z. John Zhang (2014), "Assessing Box Office Performance Using Movie Scripts: A Kernel-based Approach," *IEEE Transactions on Knowledge and Data Engineering*, 26, 2639-2648.

[35] Goeree, M. (2008): "Limited Information and Advertising in the U.S. Personal Computer Industry," *Econometrica*, 76, 1017-1074.

[36] Hausman, J. A. (1978): "Specification Tests in Econometrics," *Econometrica*, 46, 1251-1271.

[37] Heckman, J. (1979): “Sample Selection Bias as a Specification Error,” *Econometrica*, 47, 153-161.

[38] Jones, J. M. and C. J. Ritz (1991): “Incorporating distribution into new product diffusion models,” *International Journal of Research in Marketing*, 8, 91-112.

[39] Krider, R. and C. B. Weinberg (1998): “Competitive Dynamics and the Introduction of New Products: The Motion Picture Timing Game,” *Journal of Marketing Research*, 35, 1-15.

[40] Li, S., J. Mazur, Y. Park, J. Roberts, A. Sweeting and J. Zhang (2018): “Endogenous and Selective Service Choices After Airline Mergers,” working paper.

[41] Lehmann, D. and C. Weinberg (2000): “Sales through Sequential Distribution Channels: An Application to Movies and Videos,” *Journal of Marketing*, 70, 74-89.

[42] Liu, Y (2006): “Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue,” *Journal of Marketing*, 64, 18-33.

[43] Luan, Y. J. and K. Sudhir (2010): “Forecasting Marketing Mix Responsiveness for New Products”, *Journal of Marketing Research*, 47, 444-457.

[44] Moul C. C. (2007): “Measuring Word of Mouth’s Impact on Theatrical Movie Admissions,” *Journal of Economics and Management Strategy*, 16, 859-892.

[45] Murry, C. (2017): “Advertising in Vertical Relationships: An Equilibrium Model of the Automobile Industry,” working paper.

[46] Natividad, G. (2013): “Financial Slack, Strategy, and Competition in Movie Distribution,” *Organization Science*, 24, 846–864.

[47] Orbach, B. Y. (2004), “Antitrust and Pricing in the Motion Picture Industry,” *Yale Journal on Regulation*, 21, 317–367.

[48] Orbach, B. Y. and L. Einav (2007), “Uniform prices for differentiated goods: The case of the movie-theater industry,” *International Review of Law and Economics*, 27, 129–153.

[49] Orhun, A. Y., S. Venkataraman and P. K. Chintagunta (2016), “Impact of Competition on Product Decisions: Movie Choices of Exhibitors,” *Marketing Science*, 35, 73-92.

[50] Packard, G., A. Aribarg, J. Eliashberg and N. Z. Foutz (2016), "The role of network embeddedness in film success," *International Journal of Research in Marketing*, 33, 328-342.

[51] Perrigne, I., Q. Vuong and W. Yan (2019) , "Estimating Models of Entry and Differentiated Products", working paper.

[52] Radas, S. and S. M. Shugan (1998): "Seasonal Marketing and Timing New Product Intruductions," *Journal of Marketing Research*, 35, 296-315.

[53] Rao, V. R., S. A. Ravid, R. T. Gretz, J. Chen and S. Basuroy (2017): "The Impact of Advertising Content on Movie Revenues," *Marketing Letters*, 28, 341-355.

[54] Rivers, D. and Q. Vuong (1988): "Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models," *Journal of Econometrics*, 39, 347-366.

[55] Robinson, P. (1988): "Root-N-Consistent Semiparametric Regression," *Econometrica*, 56, 931-954.

[56] Seim, K. (2006): "An Empirical Model of Firm Entry With Endogenous Product Type Choices", *Rand Journal of Economics*, 37, 619-640.

[57] Shum, M. (2004): "Does Advertising Overcome Brand Loyalty? Evidence From the Breakfast-Cereal Market", *Journal of Economics and Management Strategy*, 13, 241-272.

[58] Sweeting, A. (2009): "The Strategic Timing Incentives of Commerical Radio Stations: An Empirical Analysis Using Multiple Equilibria," *Rand Journal of Economics*, 40, 710-742.

[59] Weinberg, C. B. (2005): "Profits Out of the Picture: Research Issues and Revenue Sources Beyond the North American Box Office," C. C. Moul, ed. *A Concise Handbook of Movie Industry Economics*, Cambridge University Press, 163-203.