

Information Discovery and the Long Tail of Motion Picture Content

Anuj Kumar, Michael D. Smith, and Rahul Telang[†]

{akumar1, mds, rtelang}@andrew.cmu.edu

This Version: April 2011

Acknowledgements: The authors thank Sinan Aral, Brian Kovak, and seminar participants at the Workshop on Information Systems and Economics, the Symposium on Statistical Challenges in Electronic Commerce Research, University of Utah, University of Florida, Georgia Institute of Technology, and Arizona State University for helpful comments on this research. The authors acknowledge Carnegie Mellon's iLab, the Marketing Science Institute, the Wharton Interactive Media Initiative, and the WPP-Google Marketing Research Award program for generous financial support. Rahul Telang also acknowledges financial support of the Alfred P. Sloan Foundation Industry Studies Fellowship.

[†] School of Information Systems and Management, Heinz College, Carnegie Mellon University, Pittsburgh, PA, 15213.

Information Discovery and the Long Tail of Motion Picture Content

ABSTRACT

Recent papers have shown that, in contrast to “the Long Tail” theory, movie sales remain concentrated in a small number of hits. These papers have argued that concentrated sales can be explained, in part, by heterogeneity in quality and increasing returns from social effects. Our research analyzes an additional explanation: how incomplete information may skew sales patterns. We use the movie broadcast on pay-cable channels as an exogenous shock to the availability of information, and analyze how this shock changes the resulting sales distribution.

Our data show that the pay-cable broadcast shifts the distribution of DVD sales toward “Long Tail” movies, suggesting an information spillover from the broadcast. We further develop a learning-based model of DVD demand to precisely quantify the lost DVD sales due to incomplete information. Our study contributes to the academic literatures on information provision and market outcomes, and the dynamics of long tail markets.

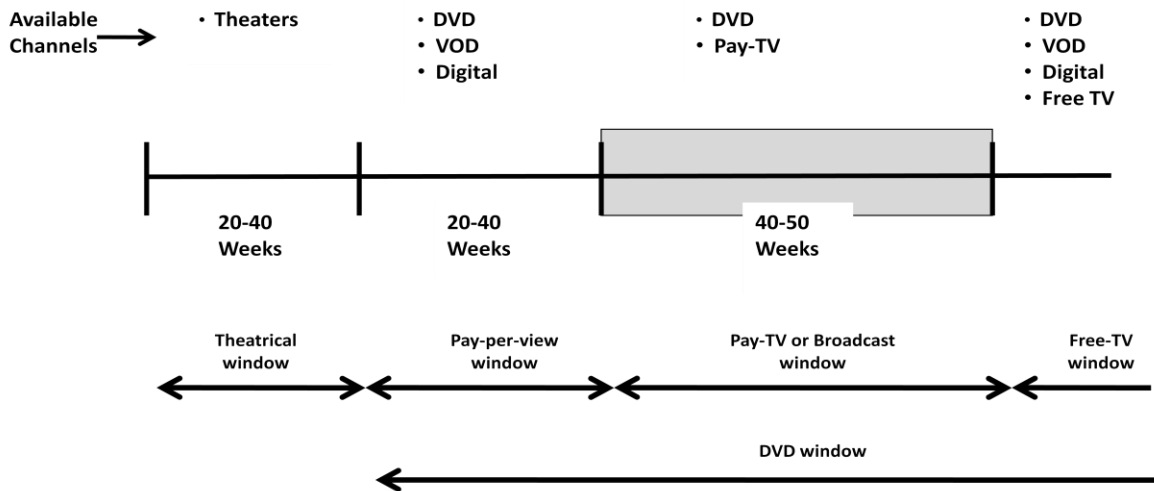
Keywords: *Incomplete information, product discovery, multichannel distribution, movie industry, learning model, Long Tail.*

1. Introduction

“The Long Tail” is a term coined by Anderson (2004) to describe an environment where digital distribution channels allow firms to promote and sell a larger variety of products than would have been possible in “physical” channels. As a result of this increased product variety available for sale, one would expect the distribution of sales to shift toward a larger variety of successful products, an outcome that has been demonstrated in the context of books (Brynjolfsson, Hu, and Smith 2003).

However, in the face of this possibility, recent studies have argued that “The Long Tail” effect is not observed in important categories of products, notably movies (e.g., Elberse and Oberholzer-Gee 2007). Explanations for the lack of a long tail outcome in these product categories include the fact that movies and music have high fixed costs of production, have externalities from the “social” nature of consumption, and feature significant heterogeneity in quality causing “good” movies to outsell “bad” movies (Elberse 2008).

Figure 1: Exclusive windows in movie lifecycle



In this paper we explore an additional explanation for the concentration in movie sales: incomplete information. We do this in the context of movies that are shown on pay-cable channels such as HBO and Showtime. Since our data allow us to fix the supply side, any shift in consumption can be attributed to demand side effects. Figure 1 shows the pay-cable release window in the context of the overall movie lifecycle. As shown, movies are first released in theaters and are available exclusively through this channel for 20 to 40 weeks. Movies are then released in DVD, Video-On-Demand (VOD), and digital channels such as iTunes. After 20 to 40 weeks in these windows, movies shift into the “pay-TV broadcast window” and are shown on pay-cable channels. Finally, movies are made available on the advertising supported cable and broadcast channels popularly called the Free-TV window.

For our study, this setting gives us a unique opportunity to examine how availability of more information impacts the skewness of the movie sales distribution. In particular, in the lifecycle described above, consumers initially become aware of movies that are distributed exclusively through local physical theaters. Because these theaters can only show a small number of movies at any given time, studios have incentives to only promote a small number of movies, and consumers are likely to become aware of only a small number of popular movies. However, by the time movies enter the pay-cable “broadcast window,” channel operators are able to show a much larger variety of movies to their wide subscriber base. Moreover, unlike the prior movie distribution windows that are all based on a per-movie fee, after paying a monthly subscription fee pay-cable subscriber can sample these movies with no additional marginal cost. This may result in consumers viewing and becoming aware of movies they otherwise would not have considered for purchase which in turn may affect the skewness of the movie sale distribution

In addition, the broadcast window has a number of important empirical characteristics that allow us to use the movie broadcast as an exogenous shock to the amount of information consumers have about movies and then analyze how this information changes the sales distribution patterns: First, the contracts between studios and pay-cable channels are such that all movies produced by a particular studio (both highly successful and less successful) are licensed to the cable channel. Second, the cable channel’s incentives are such that they broadcast all of the movies available through their license. Third, the institutional details of the movie industry are such that there is significant exogenous variation between the time a movie starts showing in theaters and the time it appears on the pay-cable channel. Finally, the timing of broadcast is such that we can isolate the effect of the broadcast (“broadcast effect”) from other changes in distribution that occurs during the broadcast window. We motivate each of these characteristics in more detail below.

The role of information in product sale is an important question for a variety of reasons. First, incomplete information represents a welfare loss to consumers who, were they fully informed, would prefer to watch a less popular movie more aligned to their taste than the more popular movie they were aware of. Second, incomplete information about products affects the product variety available in the market. Incomplete information may tilt investment towards products with mass-market appeal, rather than niche products, and thus support only a limited talent base of artists. Third, incomplete information may result in a loss of potential revenue to producers. Thus, mechanisms to provide information to consumers are of great interests to both the academics scholars and industry managers.

Our research also answers important questions on how different distribution channels interact. Managing diverse channels is a significant challenge to firms as the introduction of a product in one channel may cannibalize products in other channels leading to significant contractual challenges for media companies, and our paper sheds light on how movie broadcasting affects DVD sales and DVD rental.

To test how incomplete information affects the skewness of movie sales in the broadcast window, we identified a sample of 314 movies that were shown on the four major movie pay cable channels (HBO, Cinemax, Starz and Showtime) between January 2008 and March 2010. For these movies, we collected weekly DVD sales from release to at least 14 weeks into the broadcast window. We find that the first month DVD sales distribution for our sample of movies is almost identical to its box-office sales distribution, and that both are highly skewed. In our data the top 10% of movies make up 48% of theatrical sales and DVD sales prior to the pay-cable broadcast. A number which is identical to that reported by Elberse (2008). Moreover, the skewness of movie sales is nearly constant up to the start of the broadcast window.

However, immediately after the pay-cable broadcast, we find a disproportionately large increase in DVD sales for less popular movies relative to more popular ones. This result in a large and statistically significant reduction in the skewness of the DVD sales distribution immediately after movies enter the broadcast window, such that in the months after broadcast, the top 10% of movies make up only 35% of DVD sales.

Our empirical analysis indicates the following: (1) DVD sales for movies in our sample increase during the broadcast window, (2) the increase in DVD sales persists over the entire period of the broadcast window, (3) DVD sales for less popular and less known studio's movies increase more than the popular and major studio movies.

Based on these results, we build a movie discovery model and conduct a structural analysis to test the movie discovery explanation for the increase in movies' DVD sales. We specify a learning based function for the probability of movie discovery based on sales in previous weeks. During the broadcast window, we allow for an additional discovery parameter in the probability function to capture awareness created due to the movie broadcast. From this discovery function, we derive the weekly market demand for DVDs and estimate it on our data. We find a positive and significant estimate for the discovery parameter, indicating that the movie broadcast leads to increased discovery. We also separately estimate the parameters of the movie discovery function for the movies of major studios and of smaller studios, finding a higher discovery parameter for movies released by smaller studios.

The primary motivation for the movie discovery model is to estimate latent consumer awareness about movies at different times after the DVD release, and thus to compute the counterfactual DVD sales which would have occurred if consumers were fully aware of movie quality. At the time of broadcast window we find that 89% of potential consumers have discovered movies in the upper quartile of theatrical sales, but only 77% of potential customers have discovered movies in the lower quartile. This suggests that at the time of the broadcast window there is very little additional scope for discovery of movies which were successful in the theater; whereas, less successful movies have a large scope of discovery during the pay-cable broadcast window, leading to a proportionately higher increase in DVD sales.

We believe our paper makes several contributions. First, it provides a non-obvious result for how different channels of movie distribution interact, namely how movie distribution in one channel (TV broadcast) increase its consumption in another contemporary channel (DVD sales). Second, our setting allows for clean identification of the role incomplete information plays in the movie sales distribution and thus contributes to the growing literature on impact of information provision on market outcomes. Third, our paper highlights that how television broadcasts lead to discovery of lesser-known movies, and hence increases the sales in the tail of the movie sales distribution. Finally, our research has implications for positioning of the movie broadcast within the overall product lifecycle and for exploiting the “Long Tail” of movie demand given the large number of new digital distribution channels available to movie studios.

2. Literature Review

Our paper draws on three main literatures: the marketing literature on predicting sales patterns in the movie industry, the literature on the long tail phenomenon, and the marketing and economics literature on consumer search and the impact of information provision on market outcomes.

First, there is a rich academic literature on film marketing. Most of the academic work on the movie industry has analyzed issues around the theatrical release of the movie (e.g., Sawhney and Eliashberg 1996, Elberse and Eliashberg 2003, Sawhney and Eliashberg 1994, Ainslie et al. 2005, Krider and Weinberg 1998). In the past decade the focus of this literature has shifted to include the DVD sales channel, as revenue from DVD sales has grown to be about 46% of the total revenue (Epstein 2005, p. 20). This shift toward DVD revenue has led to a series of papers analyzing whether the DVD channel cannibalizes theatrical sales, and the optimal release time for a movie on the DVD channel (e.g., Lehman and Weinberg 2000, Luan and Sudhir 2005, Prasad et al 2004). However, we are only aware of one other paper that has analyzed the impact of the television broadcast of movies on DVD sales (Telang and Smith 2009), and this paper did not analyze information discovery or differentiate between the impact on head and tail movies.

Our paper also relates to the growing academic literature on the “Long Tail” phenomenon, a term coined by Anderson (2004) to describe how the increased stocking capacity of Internet retailers may allow niche products make up a larger share of total sales than they would in a brick-and-mortar environment. This literature has looked at a variety of issues, including how increased product variety impacts consumer surplus in books (Brynjolfsson, Hu, and Smith 2004), the demand-side and supply-side drivers of the long tail phenomenon (Brynjolfsson, Hu, and Smith 2006), the impact of niche sellers on online markets (Bailey et al. 2008), and the impact of recommender systems on the long tail (Fleder and Hosanagar 2009).¹ In

¹ See Brynjolfsson, Hu, and Smith (2010) for a review of this literature.

this literature, our work is most closely related to Elberse and Oberholzer-Gee (2007) who examine the sales concentration in the home video market from 2000 to 2005. They find that; (1) a smaller proportion of titles account for the bulk of sales over time, and (2) the number of non-selling titles has increased over time. They find some evidence of a long tail effect in that the numbers of titles that sell only few copies increase two fold during their study period, but their main finding is that the Long Tail concept is poorly suited to the characteristics of the motion picture industry. In this regard, our research identifies one potential explanation for this phenomenon: incomplete information about movies in consumers, and discusses ways in which new technology channels may reduce this lack of information.

In addition to increased product variety, Internet markets provide consumers with search tools, browsing tools, and recommendation system, and these tools may lower consumer search costs and further increase sales in long tail of the sales distribution. Several papers in the literature have developed predictions on how search cost can affect price, price dispersion, product entry, and product variety (e.g., Anderson and Renault 1999, Bakos 1999, Cachon et al. 2008, Brynjolfsson and Smith 2000, Brown and Goolsbee 2002, Hann, Clemons, and Hitt 2003). However, there are few studies analyzing how a reduction in search cost affects the concentration of product sales (a notable exception is Brynjolfsson, Hu, and Simester 2010).

Within this literature our research is closely related to the literature on search and experience goods. Nelson (1970) defines “search goods” as goods whose quality can be determined by consumers prior to consumption, and “experience goods” as goods whose quality can be completely evaluated only after consumption. Nelson (1974) further finds that advertisers provide direct information for search goods but provide indirect information or simply brand advertising for experience goods. He shows that this difference in the character of information leads to higher advertising for experience goods.

This is important for our study because movies are a classic experience good: consumers can fully evaluate their true quality only after consumption. Movie studios’ advertising therefore focuses on information about brand, director power and star power as signals of quality. Consumers typically rely on recommendations from friends and peers, and trailers to gather information about movie quality. As a result, movies from major studios, movies with early commercial success, and movies with more prominent directors and actors/actresses are advertised more and in turn get more word of mouth recommendations. These factors may drive the high concentration in movie sales that are seen in both the box-office and DVD release periods. However, this paper argues that the pay-cable broadcast window provides a new channel for “advertising” the quality of a movie by allowing pay-cable channels subscribers to evaluate movies more accurately and at a lower search cost than may have been possible in their other release windows.

Finally, our paper contributes to a growing literature analyzing the impact of information provision on market outcomes. In markets with a large number of products whose quality is difficult to determine ex-ante, consumers face a problem of incomplete information about their choice set. Goeree (2008) shows that consumers may be less than fully informed about the set of available products (in this case personal computers) because of the rapid pace of technological change. In cases when consumers are uncertain about quality, strong reputation of existing products increases demand for new products sold under the same brand name (forward spillover), and the high-quality new product can improve the brand image and thus boost the sales of the existing products (backward spillover) (Choi 1998, Cabral 2000).

Within this literature, our paper is most closely related to Hendrick and Sorensen (2009), who find a substantial and persistent increase in sales of an artist’s catalog albums due to discovery during the release of an artist’s follow-on albums. Our work extends their results to the context of movies, uses a novel natural experiment representing an exogenous shock to the availability of information about movies, and focuses on the impact of this increased information on sales of long tail products relative to “head” products.

3. Data

Our data include all movies shown on the four major U.S. pay-cable networks — HBO, Cinemax, Starz, and Showtime — between January 2008 and March 2010, comprising 314 movies in all. These channels account for the vast majority of pay-cable movie broadcasts during our study period. Among these movies, 250 are from the seven “major” Hollywood studios (Warner, Lionsgate, Sony, Paramount, Disney, Universal, and Fox) and the remaining 64 are from smaller studios. For each movie, we collected data on weekly U.S. DVD sales, the broadcast window start date, broadcast dates within that window, and the studio, genre, and box-office sales of the movie. We report summary statistics for these data in Table 1.

Table 1: Summary Statistics

| | N | Mean | Std. Dev. | Percentile | | |
|--|-----|-------------|-------------|-------------------|----------|-------------------|
| | | | | 0.10 | 0.50 | 0.90 |
| Time between DVD release and broadcast | 314 | 31.06 weeks | 13.09 weeks | 20 weeks (min 10) | 27 weeks | 55 weeks (max 96) |
| Theatrical revenue (Million US \$) | 314 | 37.32 | 55.36 | 0.06 | 15.54 | 102.36 |
| Total DVD sales | 314 | 933,639 | 1,465,847 | 9,424 | 390,905 | 2,501,037 |
| DVD sales in broadcast window | 314 | 175,583 | 287,598 | 986 | 58,506 | 529,838 |
| % of total DVD sales in first 4 weeks | 314 | 54.52 | 15.00 | 33.20 | 57.03 | 71.57 |
| % of total DVD sales in broadcast window | 314 | 17.38 | 13.39 | 3.16 | 13.91 | 35.08 |

These summary statistics show that the DVD release dates and broadcast window dates for movies in our sample are spread out fairly evenly over the calendar year. There is also a large variation between the time of the DVD release and the start of the broadcast window in our sample of movies: the 10th and 90th percentile figures in our data are 20 and 55 weeks respectively. This is an important factor that we will utilize in our identification strategy later.

Table 1 also shows that DVD sales are highly skewed across titles and over time. Across titles, the mean sales in the broadcast window are 175,583 and the median is 58,506. Across time, we see that on average 55% of DVD sales occur in the first four weeks after release. Still, 17% of total DVD sales occur in the broadcast window, highlighting the fact that an economically significant number of sales occur during the broadcast window.

One may worry that pay-cable channels pick-and-choose which movies to show on their networks (presumably choosing only the most popular movies) leading to selection problems. However, based on reporting in the trade press, this seems unlikely. In the United States, the major pay-cable channels negotiate multiyear deals with studios, called output deals, wherein the pay-cable channel pays the movie studio a fixed sum to get the exclusive broadcast rights for all movies that the studio releases during the negotiated period. Consider the following quotes from the industry magazines *Broadcast & Cable* and *Variety*:

“Universal Pictures and HBO renewed their domestic output deal, extending it midway into the next decade. The pact gives HBO the right to program all movies from Universal and its specialty labels Focus Features and Rogue Pictures both on TV and online.” (Becker 2007)

“You're buying stuff that, in many cases, hasn't even been thought of yet ... When you do an output deal, you're betting on the studio. You're saying, ‘I want all the films that this studio releases over the next however many years.’” (Frankel 2009)

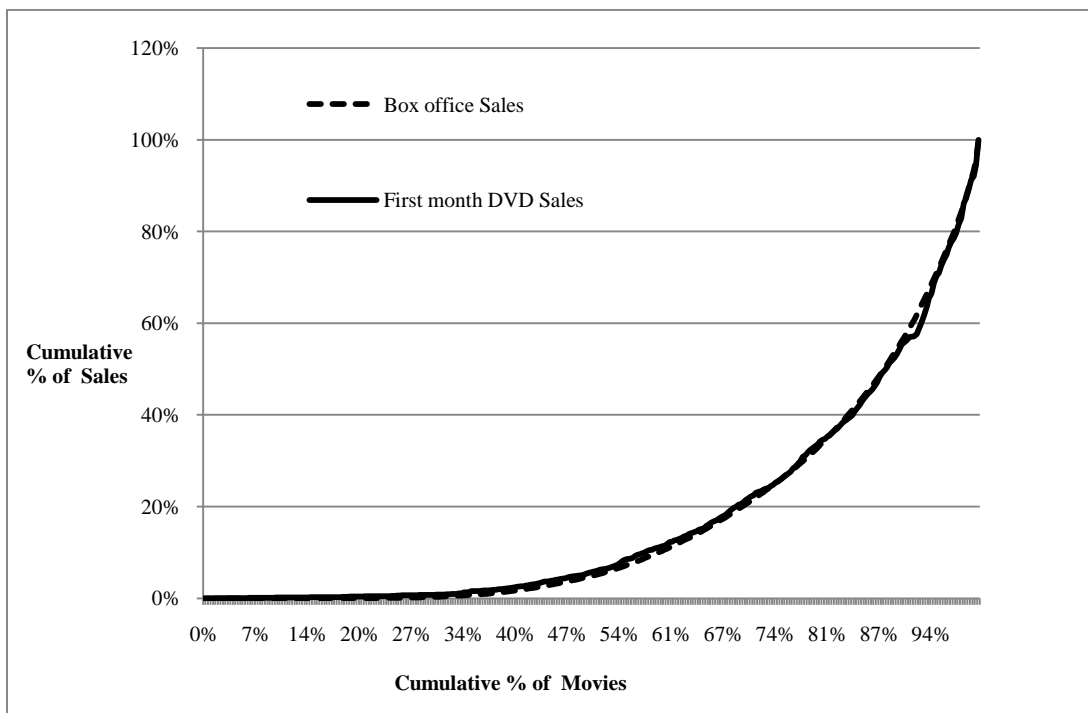
In summary, (1) output deals are negotiated for all movies released by a particular studio (as opposed to on a movie-by-movie basis), and (2) these deals are negotiated 5-10 years in advance, and thus the box office performance of the included movies is not known at the time the deal is made. We confirmed these contract characteristics through interviews with two industry representatives who have detailed knowledge of output deal negotiation. These interviews reveal that only documentaries and movies of some small studios aren't covered in these output deals. Thus, essentially all mainstream movies are available for broadcast on pay-cable channels, whether it is successful in the box-office or not.

The next challenge is whether all movies covered by these licensing deals are ultimately shown on the cable channel, and whether the cable channels are strategic in scheduling the movie broadcast (for exam-

ple by broadcasting successful movies earlier than less successful movies). Based on our discussions with the same two industry representatives, we find (1) that the output deal gives the cable channel the right to broadcast movies starting from a set time (typically 9 months to a year) after the month of the theatrical release, (2) that this lag is the same for all movies covered by the contract, (3) that the cable channels typically show every movie covered by the contract, and finally (4) that pay-cable channels begin showing the movies immediately after it becomes available under the contact to maximize its exposure during the license window period.

As a result, we expect the lag between the theatrical release and the start of the pay-cable broadcast window to be similar for successful and less successful movies within a contract. Thus, for our sample of broadcast movies from different contracts, we expect that movies will enter the broadcast window in a way that is not systematically related to box-office success. This is an important identification requirement in our econometric specification and we will explicitly test this identification requirement below.

Figure 2: Distribution of box office sales and first month DVD sales for broadcast movies



3.1. DVD Sales Distribution and its Evolution in the Broadcast Window

Next, we compare the skewness of box office sales to the skewness of DVD sales — and analyze how this skewness changes between the initial DVD release and the pay-cable broadcast window. To do this, we take the Lorenz curve for box-office sales of the movies in our sample, and *keeping the same rank order*

of increasing box-office sales for movies, we construct the Lorenz curve for DVD sales in the first month after release. We note that by keeping the rank ordering of movies in our sample the same for both Lorenz curves, we are specifically testing for the similarity between these two curves at the movie level as opposed to at an aggregate level. As is evident from Figure 2, the sales distribution for DVD sales in the first month after release is almost exactly the same as the distribution for box-office sales, and both curves have a Gini Coefficient of 0.67.

As noted earlier, movies enter the broadcast window 20-50 weeks after their DVD release. We now examine how the distribution of DVD sales changes from immediately before the broadcast window to immediately after the broadcast window. Using the same rank order of increasing box-office sales for our sample of movies, in Figure 3 we construct the Lorenz Curve for DVD sales of movies one month before the broadcast window starts and one month immediately after the broadcast window.

Figure 3: Distribution of DVD sales immediately before and after the broadcast window

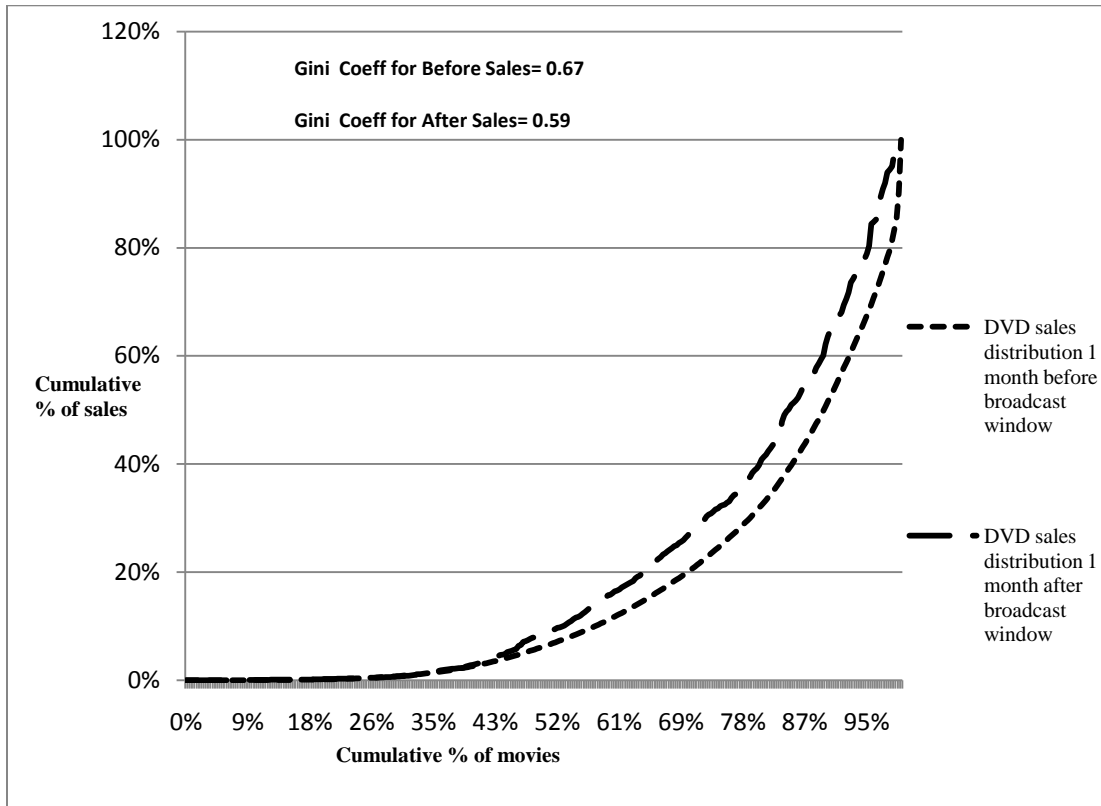


Figure 3 shows that the proportion of DVD sales for low and moderately selling movies increases considerably immediately after the broadcast window. This is reflected in a reduction in the concentration of the DVD sales distribution for our sample of movies after the start of the broadcast window. The Gini coefficient for the DVD sales distribution reduces from 0.67 before the broadcast window to 0.59 in the

broadcast window.² This is also reflected in a shift away from sales in the top movies. Prior to the pay-cable broadcast the top 10% of movies accounted for 48% of the DVD sales (which is exactly the same as the skewness reported in Elberse 2008 for Quickflix movie rentals), whereas after the pay cable broadcast the top 10% of movies account for only 35% of DVD sales.

Thus the raw data suggests that, in the broadcast window, moderate and less popular movies get a higher increase in DVD sales than more popular movies do. We test this finding more rigorously in the following sections.

4. Empirical Model and Results

4.1. Blackout and Broadcast Effects

We now formally analyze the change in DVD sales of a movie due to its broadcast on pay-cable channel. To do this, we must first isolate changes in DVD sales that occur due to factors other than the movie broadcast during the broadcast window. Specifically, we note that when a movie is shown on pay-cable networks, licensing contracts require that the studio provide the cable channel with an “exclusive broadcast window.” Historically, this exclusive window meant that the studio cannot license the movie to show on another competing broadcast channel. With the advent of new distribution channels such as cable Video-On-Demand (VOD) and Electronic channels such as iTunes, this exclusive license has been interpreted to include these two channels.

Because of this, when the broadcast license begins, the studio is required to remove the licensed movie from sale on both the cable VOD and Electronic channels. However, importantly for our study, the movie studio is not required to remove the movie from the DVD channel, in part because once the physical goods have entered the distribution channel it would be nearly impossible to do so. In view of this setup, there are potentially two distinct effects of the movie broadcast window on DVD sales:

- (1) *Blackout effect*: A consumer who wants to purchase a movie through a cable VOD or Electronic channel can't do so (due to the blackout of the movie on these channels) and may purchase the movie on DVD instead.
- (2) *Broadcast effect*: Broadcasting a movie can have two opposite effects on DVD sales:
 - (a) *Information spillover effect*: Broadcasting a movie on a pay-cable channel enables subscribers to sample the movies at a low cost (given the fixed monthly cable subscription

² Note that the same value of the Gini coefficient for the DVD sales distribution before the broadcast window and after DVD release indicates no change in the DVD sales distribution from the time of DVD release to the beginning of broadcast window.

fee) and thus evaluate the quality of the movie. These informed subscribers may generate word of mouth for the movies, or they may purchase the movie if they like it. In either case the result will be an increase in DVD sales.

- (b) *Cannibalization effect:* Broadcasting the movie on pay-cable channels may also cause consumers who would have otherwise purchased the DVD to consume it on the pay-cable channel instead. This cannibalization effect could be more severe given the prevalence of Digital Video Recorders and other technologies that allow customers to digitally record the movie for later playback.

To separately identify the “blackout effect” from the (net) “broadcast effect” we need exogenous variation between the time the blackout effect starts and the time the movie is shown on the cable channel. To do this, we exploit the fact that the output deal contracts specify that the blackout period starts on the first day of the month in which the movie becomes available to the cable channel, and that cable channels typically broadcast newly available movie on the first, second, third or fourth weekend of the month. For example, in March 2011 the movies “Robin Hood,” “MacGruber,” “Cop Out,” and “Just Wright” premiered on HBO. These movies were all removed from iTunes and VOD channels on March 1 and were first broadcast on HBO on March 5, 12, 19, and 26 respectively. Our estimations use the natural variation between the start of the blackout period and the first broadcast to isolate these two effects.³

Given this environment, we separately estimate the “broadcast period” and “blackout period” treatment effects on DVD sales of movies with the following specification:⁴

$$\text{Log}(S_{it}) = \alpha_i + \beta_{black} \times D_{it}^{black} + \beta_{broad} \times D_{it}^{broad} + \sum_t \delta_t \times D_t + \sum_{wk} \delta_{wk} \times D_{wk} + \varepsilon_{it} \quad (1)$$

where $\text{Log}(S_{it})$ is the log of DVD sales in week t for movie i , and t is measured in weeks since release on the DVD channel. On the right-hand side of specification (1), α_i is the movie fixed effect, D_t are indicator variables for t in weeks since the DVD release, and D_{wk} are indicator variables for calendar weeks. Our variables of interest are D^{black} and D^{broad} , indicator variables for when movies are blacked out from the VOD and Electronic channels, and for the period after movie is initially shown inside the broadcast window respectively. Therefore, β_{black} represents the impact of the blackout effect on DVD sales over the entire broadcast window, and β_{broad} represents the additional impact of the broadcast effect on DVD sales. The calendar week indicator variables capture weekly shocks to DVD sales, for example competitive effects from the entry of other movies or seasonal effects from holiday buying. Indicator variables for each

³ In Table A2 Appendix A, we show that in our sample there is no systematic relationship between the lag between the broadcast from blackout period and the box-office sales of the movies.

⁴ Our approach follows Hendrick and Sorensen (2010), who apply an empirical approach from the literature on treatment effects (e.g., Wooldridge 2002).

week since release (D_t) captures a flexible (non-parametric) form of average decay path of DVD sales for all movies.

Intuitively, after controlling for other variations (movie fixed effects, weekly calendar dummies and sales decay over time) in equation (1), β_{black} and β_{broad} represent the average treatment effect (ATE) on the treated movie for both the blackout period and the broadcast period. Put another way, the movie which enters the window at time t is the treated movie and the other movies which have not entered the window by time t act as control. Therefore, the $ATE_t = \log(S_{it}^T) - \log(S_{it}^O)$, where $\log(S_{it}^T)$ is the log of DVD sales for movie i in week t in the treatment window, and $\log(S_{it}^O)$ is the log of DVD sales for movie i in week t not in the treatment window. Since movie i cannot simultaneously be in and out of the window we use the movies that have not yet entered the window as the control group against which we compare the sales of the movies which have entered the window. Therefore, for a movie that enters the broadcast window at week t , the counterfactual sales can be inferred from the sales of all the movies that have not entered the broadcast window at week t .

In addition to having exogenous variation between the beginning of the blackout period and the first broadcast, here are two requirements for clean identification of the treatment effect in specification (1). First, there should be variation in the time between the DVD release and the broadcast window to ensure that we have sufficient untreated movies at any given t to estimate the treatment effect. Second, the movie entry time into the broadcast window should not be systematically related to its characteristics (for example, commercial success). In short, the entry of movies in the broadcast window should be more or less random.

With respect to the first requirement, the summary statistics in Table 1 show that the average movie enters the broadcast window 27 weeks after the DVD release, with the earliest 10% of movies entering before 20 weeks and the latest 10% of movies entering after 55 weeks. There are two main reasons for this variation. First, movies in our sample are from several different studio output deals, and thus several different minimum time periods between theatrical or DVD release and the start of the broadcast window. Second, since the broadcast window for movies begins on first day of a month but the movies could be released in theaters (DVD) in any week of the month, we see an additional four weeks of variation in the time between the theatrical (DVD) release and the broadcast window even for movies from the same output deal.

As far as movie entry in the window is concerned in Section 3 we showed that output deals between studios and cable networks are done for all movies to be released in next 5-10 years (well before many movies are even conceived). Thus, contractually there is no reason to expect a relationship between the commercial success of a movie and its time of entry into the broadcast window. We now explicitly test whether the movie entry time in the broadcast window is correlated with any movie characteristics. We

report the results in the Appendix A and find no systematic relationship between movie characteristics (box office success, Genre, small vs big studio and so on) and the release lag before broadcast. Therefore, in our sample of movies, we find no systematic variation in movie entry into the broadcast window. This satisfies a key identification requirement and highlights an important methodological contribution of our research.

To estimate the treatment effect, we use the total broadcast window of 30 weeks. We do this because we find that weekly DVD sales are almost negligible after this period. The stochastic error term in specification (1) is assumed to be heteroskedastic across movies (some movies sales are more volatile than others) and auto-correlated within movies (random shocks to sales are persistent over time). Note that the coefficients of indicator variables for each week since release (D_t) capture an average non-parametric decay path of DVD sales for movies in our sample. The difference in the scale of DVD sales for movies is captured by the movie fixed effect but the difference in the DVD sales decay path becomes part of the error term. Table 2 shows the results of estimating specification (1) on our data.⁵

Table 2: Estimates for DVD sales

| Dependent Variable: Log(DVD sales) | Coefficient Estimates (Std. errors in parenthesis) |
|------------------------------------|---|
| Blackout period dummy | 0.096 ^{***} (0.032) |
| Broadcast period dummy | 0.101 ^{***} (0.050) |
| N | 17194 (314 movies) |
| R sq | 0.946 |
| Movie fixed effect | yes |
| Cluster correction at movie level | yes |
| Heteroskedasticity correction | yes |

Note - ***, **, * = statistically significant at the 1%, 5% and 10% levels (two-sided test), respectively.

The estimate for the blackout period dummy and the broadcast period dummy are positive and significant. The estimates suggest that DVD sales increase by 10% due to the blackout effect and by an additional 12% due to the broadcast effect of the movie.⁶ These estimates translate into, on average, an additional 464 DVD sales per week for each movie due to the broadcast window. For the 30 week broadcast window and for our sample of 314 movies, this translates into additional sales of approximately 4.4 million DVDs

⁵ Note that the difference in sales decay rates for individual movies from the average becomes part of the error term in specification (1). If these differences in decay rates are systematically correlated to their time of entry in broadcast window, it may bias the treatment effect estimates in specification (1). In Appendix B, we test for this and find that our results are robust to this possibility.

⁶ To save space we have not included the coefficient estimates for time indicators in Table 2, but these coefficients are consistent with expectations: DVD sales decline with time.

and additional revenue of approximately \$66 million (assuming an average DVD sales price of \$15). For readability, we have not reported the estimates on weekly dummies and movie release time dummy.

Note that the treatment effect in specification (1) is computed by comparing decay rates of treated movies with untreated movies. It may be a possible that our results are due to specific decay patterns for movies in our sample rather than the broadcast window effect such that we would get the similar treatment effects by running specification (1) even in the pre-broadcast window period for our sample of movies. We test for this possibility in Appendix-B by randomly inducing a placebo treatment of broadcast in the pre-broadcast window period for our sample of movies. We do not find any statistically significant coefficients for the placebo treatments, suggesting that our results are not driven by movie-specific decay patterns.

Note that by taking the log of DVD sales, we are measuring the treatment effect in proportional terms. The reason for this is the highly skewed distribution of DVD sales for our sample of movies, and thus the average treatment effect is likely to be nonlinear: the magnitude of the increase in absolute DVD sales from the television broadcast of a popular movie is likely to be much higher than it is for less popular movies. The average treatment effect in proportional term is likely to capture some of these nonlinearities. However, this may bias our treatment effect estimates upwards since the proportional treatment effect is likely to be higher for less popular movies, and there are more of them in our sample. In Appendix-C (Table C2), we compute a linear treatment effect for different subsamples of movies with less skewed DVD sales to show that our results are not merely due to computing the proportional treatment effect. We further compute in Appendix-C (Table C1) the proportionate treatment effect for popular movies that enter the broadcast window later (i.e. with low values of DVD sales at the broadcast window) and still find a small and insignificant proportional treatment effect, indicating that our results are not due to computing proportional treatment effect for movies with small DVD sales.

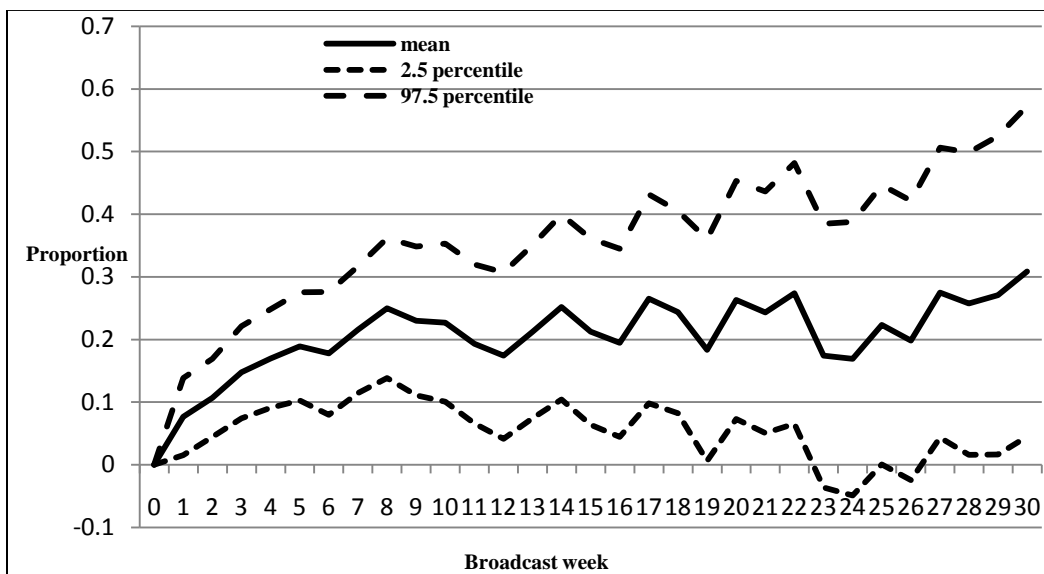
It is also notable that the treatment effect varies with time within the broadcast window. Specifically, we can examine whether the increase in DVD sales we observe is due to an inter-temporal demand shift for movies: consumers who would have otherwise purchased the DVD shifted their purchase ahead in time due to awareness created by its TV broadcast. We examine this possibility by including weekly dummies for each week in the broadcast window, using the following regression model (2):

$$\text{Log}(S_{it}) = \alpha_i + \sum_{brwk} \delta_{brwk} \times D_{brwk} + \sum_t \delta_t \times D_t + \sum_{wk} \delta_{wk} \times D_{wk} + \varepsilon_{it} \quad (2)$$

where D_{brwk} are indicator variables for each week in the 30 week broadcast window and all other variables are the same as in (1). To save space, in Figure 4 we report the coefficient estimates for (2) in a graphical form. This figure shows that all weekly coefficients are positive and, with a few exceptions, statistically

significant. If the spillover effect represented a transitory demand shift, we would expect the coefficient estimates for the broadcast week dummies to decline and eventually become negative and significant.

Figure 4: Weekly broadcast effect in broadcast window



Next we examine what happens to DVD rentals during the broadcast window. To do this, we use data from RENTRAK⁷ for weekly DVD rentals for the movies in our sample. We could only obtain the complete weekly DVD rentals data for 194 movies out of the total sample of 314 movies.⁸ For this analysis, we use specification (1) with the log of weekly DVD rentals as the dependent variable.

Table 3: Estimates for DVD rentals

| Dependent Variable: Log(DVD rentals) | Coefficient Estimates (Std. errors in parenthesis) | |
|--------------------------------------|---|---------|
| Blackout period dummy | 0.085 ^{***} | (0.024) |
| Broadcast period dummy | -0.031 | (0.026) |
| N | 8816 | |
| R sq | 0.957 | |
| Movie fixed effect | Yes | |
| Cluster correction at movie level | Yes | |
| Heteroskedasticity correction | Yes | |

Note - ***, **, * = statistically significant at the 1%, 5% and 10% levels (two-sided test), respectively

⁷ According to company materials, RENTRAK is the world's largest reporter of DVD sales and rentals, providing detailed data on more than one billion home video rental transactions each year.

⁸ We found that RENTRAK does not track some movies from some small studios or less commercially successful movies, and thus our sample of 194 movies with sales rental data covers the more commercially successful movies in our larger sample of 314 movies.

Table 3 reports the coefficient estimates. A positive and significant coefficient estimate for the blackout period indicates that DVD rentals increase by 9% due to substitution from the cable VOD and Electronic EST channels to the DVD rentals. However, an insignificant coefficient estimate for the broadcast period indicates that there is (statistically) no additional increase in DVD rentals from the broadcast period. This finding makes sense for several reasons. First, note that the price of cable VOD and electronic rentals (such as iTunes rentals) are approximately \$3-\$4, which is close to the price of DVD rentals (\$1-\$2). Therefore, we would expect cable VOD and electronic consumers to substitute their consumption with DVD rentals. Second, the lack of an increase in DVD rentals due to movie broadcast may indicate that the DVD rental channel is a close substitute for watching the movie on a cable channel, and thus broadcasting a movie on a cable channel may cannibalize DVD rentals.

Recognizing that the error terms for the DVD sales and rentals regressions may be correlated, we used the Seemingly Unrelated Regression (SUR) model to jointly estimate specification (1) for DVD sales and DVD rentals for a sample of 194 common movies. These SUR estimates are qualitatively similar to the blackout and broadcast period estimates reported above and are available from the authors upon request.

4.2. Variation in Treatment Effect across Movies

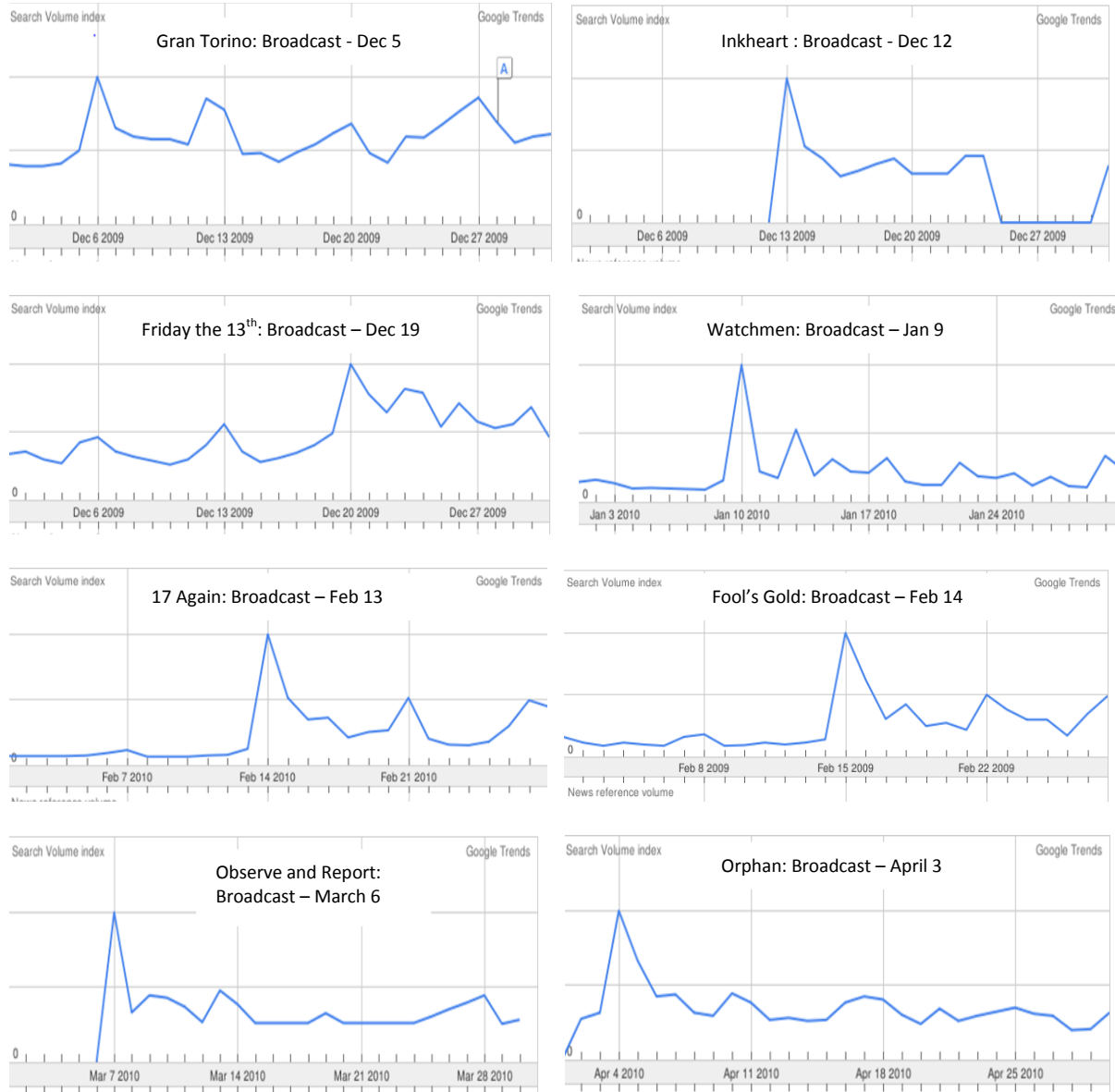
In the previous section we found a positive and significant broadcast effect over and above the blackout effect. As noted earlier, the broadcast effect is the net of the information spillover effect and the cannibalization effect. Hence, a net positive broadcast effect indicates a positive information spillover effect during the broadcast period. Moreover, we observe that the distribution of DVD sales changes from before the broadcast window to just after the broadcast window (Figure 3). Specifically, we observe that DVD sales increase more for low and moderately popular movies than for hit movies, and this change is not explained by the blackout effect. In this section we will attempt to show that it can be explained with an information spillover effect. Specifically, we analyze whether the observed increase in DVD sales after the pay-cable broadcast is caused by information spillover as consumers become better informed about the true quality of movies of which they were previously unaware (or under-aware).

A preliminary test of an information spillover effect is to see whether the TV broadcast raises general consumer awareness about movies. One way is to see whether the consumer online search about movies increases during TV broadcast. To analyze this possibility, we collected data on the Google Trends search volume index for the movies in our sample during its broadcast window. Figure 5 reports these search trends for a representative set of movies.⁹ These trends indicate that searches for the movie name increase

⁹ Other movies in our sample show comparable results.

substantially around the date of its first broadcast on the pay-cable channel, and persist at this higher level for several weeks after broadcast.

Figure 5: Trends of movie name search on Google Trend



While this is suggestive, we can conduct a more systematic test for information spillovers by analyzing how the increase in sales varies as a function of movies with high and low box-office sales. Specifically, in Table 4 we divide our sample of movies into quartiles based on box-office sales, and use specification (1) to separately estimate the blackout and broadcast period coefficients for each quartile of movies. The mean box office sales revenue (and range) for each quartile of movies in our sample are as follows: \$110 million (> \$50 million), \$32.7 million (between \$16 million and \$50 million), \$7.3 million (between \$1 million and \$16 million), and \$0.2 million (<\$1 million) respectively.

Table 4: Differential increase in DVD sales based on movie popularity

| <i>Dependent variable : Log(DVD sales) (Std. errors in parenthesis)</i> | <i>(1) Top Quartile</i> | <i>(2) Second Quar- tile</i> | <i>(3) Third Quartile</i> | <i>(4) Bottom Quartile</i> |
|---|-----------------------------|--------------------------------------|-------------------------------|--------------------------------|
| Blackout period dummy | -0.057 (0.053) | 0.072 (0.057) | 0.052 (0.060) | 0.057 (0.088) |
| Broadcast period dummy | -0.092 (0.078) | -0.008 (0.063) | 0.134** (0.067) | 0.198*** (0.089) |
| N | 4105 | 4061 | 4221 | 4807 |
| R sq | 0.875 | 0.836 | 0.889 | 0.864 |
| Movie fixed effect | Yes | Yes | Yes | Yes |
| Cluster correction at movie level | Yes | Yes | Yes | Yes |
| Heteroskedasticity correction | Yes | Yes | Yes | Yes |

Note - ***, **, * = statistically significant at the 1%, 5% and 10% levels (two-sided test), respectively

Note that the coefficient estimates for the broadcast period are both large and statistically significant for the bottom two quartiles, while they are small and insignificant for the top two quartiles. This suggests that the observed increase in DVD sales after broadcast is disproportionately felt among less popular movies. This is consistent with an information diffusion effect: When popular movies are shown on cable television, consumers are already likely to be well informed about their quality and are thus less likely to change their purchase behavior following the broadcast; whereas for less popular movies, consumers seem to obtain information from the broadcast that changes their purchase behavior.

Table 5: Broadcast effect for movies from major studios versus other studios

| <i>Dependent variable : Log(DVD sales) (Std. errors in parenthesis)</i> | <i>(1) Minor studios</i> | <i>(2) Major studios All movies</i> | <i>(3) Major studios lower half pop- ular movies</i> |
|---|----------------------------------|---|--|
| Blackout period dummy | 0.129 (0.102) | 0.034 (0.033) | 0.027 (0.050) |
| Broadcast period dummy | 0.152*** (0.057) | 0.028 (0.043) | 0.080** (0.036) |
| N | 3768 | 13426 | 6765 |
| R sq | 0.898 | 0.941 | 0.923 |
| Movie fixed effect | Yes | yes | yes |
| Cluster correction at movie level | Yes | yes | yes |
| Heteroskedasticity correction | Yes | yes | yes |

Note: ***, **, * = statistically significant at the 1%, 5% and 10% levels (two-sided test), respectively.

Our data also contain movies from major (Paramount, Warner Brothers, Disney, Lionsgate, and Universal) and smaller studios. Major studios typically promote their movies more aggressively, and with higher

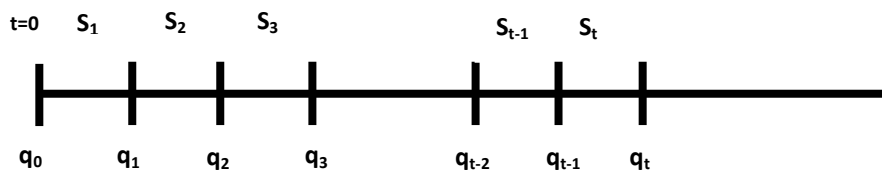
advertising budgets, than “minor” studios do, and thus consumers may have a higher likelihood of discovering movies of major studios prior to their broadcast on television. To analyze this possibility, in Table 5 we estimate specification (1) separately for movies promoted by major studios and smaller studios.

Consistent with our intuition, this regression shows a significant increase in DVD sales for the minor studios and for the less popular movies promoted by major studios. Thus, in total our results indicate a higher increase in DVD sales of the less popular movies during the broadcast period, which as we argued above is consistent with an information spillover effect.

4.3. Movie Discovery Model

In the previous sections, we saw that DVD sales of less-known movies increase proportionately more during their broadcast than DVDs of well-known movies do. We attributed this higher increase in DVDs of less-known movies to a higher proportion of consumers discovering such movies during their broadcast. In this section we specify a parametric function for the probability of movie discovery to precisely model the proportion of consumers discovering movies in a time period. We then estimate the parameters of this model on our data and use the estimated parameters to compute the lost DVD sales for movies due to incomplete information to consumers. This model is similar to the model proposed by Hedrick and Sorensen (2009).

Specifically, let $t = 0, 1, 2, 3, \dots$ represent the weeks since the release of a movie on DVD, such that $S_1, S_2, S_3, S_4, \dots$ are its DVD sales in these weeks, and $q_0, q_1, q_2, q_3, \dots$ are the proportion of potential consumers who have discovered the movie at the end of these weeks.



During the DVD release period, a consumer learns about the movie from advertisements or from consumers who are already aware of the movie. Moreover, consumers who discover the movie in week t spread word-of-mouth to other consumers in week $t+1$. Therefore, we model the probability that a consumer discovers a movie in week 1 with the following logistic learning function:¹⁰

$$P_1 = \frac{ae^{bS_1}}{(1-a)+ae^{bS_1}}$$

¹⁰ In this function, we assume that the word of mouth created for a movie in a week largely comes from the consumers who consumed it in that week. We also tried modeling word-of-mouth from cumulative sales of movies up to the week and found qualitatively similar results.

Note that if the sales of a movie in week 1 (S_1) are very small, the probability of discovery is a . Therefore, the parameter a indicates the awareness of the movie from all factors other than the current week's sales. Similarly, as S_1 gets large, the probability converges to 1. Therefore, magnitude of b determines the rate at which the sales in prior week lead to movie discovery. The probability of the purchase of a movie is the product of two probabilities: the probability that a consumer likes the movie conditional of having discovered it, and the probability that the consumer discovers the movie. Therefore, DVD sales in week 1 can be given as

$$S_1 = P_{pur|dis}^1 * q_0 * N$$

where, $P_{pur|dis}^1$ is the probability that a consumer purchases the movie conditional on having discovered it in week 1 (i.e., the consumer preference for the movie in week 1) and q_0 is the proportion of consumers who have discovered the movie at the time of release. N is the total number of potential consumers who will purchase the movie.

The proportion of aware consumer at the end of week 1 is then given by

$$q_1 = q_0 + (1 - q_0) * P_1 \Rightarrow (1 - q_1) = (1 - q_0) (1 + P_1)$$

Therefore, DVD sales in week 2 is given by

$$S_2 = P_{pur|dis}^2 * (q_1 - q_0) * N = P_{pur|dis}^2 * N * (1 - q_0) * P_1$$

The proportion of aware consumers at the end of week 2 is given by

$$q_2 = q_1 + (1 - q_1) * P_2 \Rightarrow (1 - q_2) = (1 - q_1) (1 + P_2) = \Rightarrow (1 - q_0) (1 + P_1) (1 + P_2)$$

So, we can write the DVD sales for any week $t > 2$ as

$$S_{(t-1)} = P_{pur|dis}^{(t-1)} * N * (1 - q_0) (1 + P_1) (1 + P_2) \dots (1 + P_{t-3}) * P_{t-2}$$

$$S_t = P_{pur|dis}^t * N * (1 - q_0) (1 + P_1) (1 + P_2) \dots (1 + P_{t-2}) * P_{t-1}$$

Getting rid of the unobserved N by taking the ratio of DVD sales for consecutive weeks, we obtain:

$$\frac{S_t}{S_{(t-1)}} = \frac{P_{pur|dis}^t}{P_{pur|dis}^{t-1}} \times \frac{(q_1 - q_{t-1})}{(q_{t-1} - q_{t-2})} = \frac{P_{pur|dis}^t}{P_{pur|dis}^{t-1}} \times \frac{(1 + P_{t-2}) * P_{t-1}}{P_{t-2}}$$

Consumer preferences for a movie attenuate with the age of the movie: newly released movies capture consumer's attention more than older movies do. We incorporate this by allowing the consumer's probability of purchase, conditional on discovery, to decline with age for the movies in our model. We also incorporate holiday effects in our model by assuming that the consumers' probability of purchase conditional on discovery increases in the holiday season. We specify both of these factors as follows

$$\frac{P_{pur|dis}^t}{P_{pur|dis}^{t-1}} = e^{-\delta_t + \delta_h I_h}$$

where δ_t captures the average decay of consumer preferences for a movie with time, I_h are indicator variables for holiday weeks, and δ_h capture the holiday effect. Incorporating these preferences, we get

$$\frac{S_t}{S_{(t-1)}} = \frac{(1+P_{t-2}) * P_{t-1}}{P_{t-2}} \times e^{-\delta_t + \delta_h I_h}$$

And after taking logs and simplifying, we get

$$\text{Log} \left(\frac{S_t}{S_{(t-1)}} \right) = b(S_{t-1} - S_{t-2}) + \text{Log} \frac{\{(1-a)+2a e^{bS_{t-2}}\}}{\{(1-a)+a e^{bS_{t-1}}\}} - \delta_t + \delta_h I_h$$

Inside the broadcast window ($t \geq t_{brd}$), pay-cable channels start broadcasting the movie. Therefore, during the broadcast window, consumers have additional opportunity to discover the movie via broadcast. Therefore, the discovery probability in week t inside the broadcast window is enhanced and given by

$$P_t = \frac{a e^{bS_t + c}}{(1-a) + a e^{bS_t + c}}$$

where, parameter c captures the promotional effect of the broadcast window. A positive value of c indicates that the movie broadcast leads to a higher probability of discovery.

So for any week $t > 2$, we take the following general specification (3) for the ratio of DVD sales in two consecutive weeks to the data

$$\text{Ln} \left(\frac{S_t}{S_{(t-1)}} \right) = b(S_{t-1} - S_{t-2}) + c * I_{brd1} + \text{Ln} \frac{\{(1-a)+2a e^{bS_{t-2} + c * I_{>brd2}}\}}{\{(1-a)+a e^{bS_{t-1} + c * I_{>brd1}}\}} + \delta_t + \delta_h I_h + \varepsilon \quad (3)$$

where, $I_{brd1} = 1$ if $t = t_{brd} + 1$ and 0 otherwise, $I_{>brd1} = 1$ for $t > t_{brd} + 1$ and 0 otherwise, and $I_{>brd2} = 1$ for $t > t_{brd} + 2$ and 0 otherwise.

For week $t=2$, the ratio of sales for week 2 to week 1 is given as specification (4)

$$\frac{S_2}{S_1} = e^{\delta_t + \delta_h I_h} \times \left(\frac{1-q_0}{q_0} \right) \times \frac{a e^{bS_1}}{(1-a) + a e^{bS_1}} + \varepsilon \quad (4)$$

4.4. Estimated Model Results

We estimate specification (3) on the weekly DVD sales data for our full sample of movies (for weeks 3 and higher) to get estimates for parameters δ_t , δ_h , a , b and c . We substitute these parameter values in specification (4) and then estimate it on data for our sample of movies (for weeks 1 and 2) to get estimates for parameter q_0 . We also estimate these parameters separately for popular and less popular movies in our sample. Table 6 reports the resulting coefficient estimates.

The estimates for parameter c are positive and significant, indicating that the probability of discovery increases in the broadcast window. We also find a much higher value of parameter c for movies from smaller studios as compared to major studios. This is in line with our earlier results from the reduced form estimates indicating a higher discovery for movies of smaller studios during the broadcast window.

Table 6: Estimates for movie discovery model

| <i>Non-linear least square estimates (Std. errors in parenthesis)</i> | (1) <i>For all movies</i> | (2) <i>For major studio movies</i> | (3) <i>For smaller studio movies</i> |
|---|------------------------------|---|---|
| Discovery at DVD release (q_0) | 0.147*** (0.007) | 0.217*** (0.012) | 0.078*** (0.002) |
| Learning function parameters | | | |
| a | 0.046*** (0.007) | 0.044*** (0.006) | 0.082** (0.043) |
| b | 1.683*** (0.119) | 1.620*** (0.105) | 6.878** (3.067) |
| c | 0.135*** (0.026) | 0.103*** (0.026) | 0.277*** (0.074) |
| Average preference decay (δ_t) | -0.132*** (0.008) | -0.129*** (0.008) | -0.166*** (0.008) |
| Holiday dummies (δ_h) | Yes | Yes | Yes |
| N | 17,344 | 13,556 | 3,788 |
| R sq | 0.969 | 0.964 | 0.943 |

Note - ***, **, * = statistically significant at the 1%, 5% and 10% levels (two-sided test), respectively

Specifically, estimates of q_0 suggest that overall 14.7% of potential consumers have discovered the movie at the time of their DVD release; whereas, 21.7% of potential consumers have discovered movies of major studios at their DVD release and 7.8% have discovered movies of smaller studios.¹¹ These estimates are consistent with the fact that major studios spend much more on advertising than the smaller studios do, and are consistent with our prior reduced form results. We also find a negative and significant estimate for δ_t , and positive and significant estimates for the holiday coefficients (δ_h) indicating, as expected, that preferences for movies decline with time and indicating higher DVD sales in holiday weeks.¹² Thus the parameter estimates of the movie discovery model are in line with our results from the reduced form specifications (1) and are consistent with our theory of movie discovery due to broadcast.

Next, we compute the economic significance of the estimated parameter c . Using equation (3) and the estimated parameter values, we compute the average weekly DVD sales in the broadcast window for our sample of movies with $c=0.135$ and with $c=0$, where the difference in the two DVD sales values indicates the increase in DVD sales due to movie discovery. We find that an additional 504 DVDs are sold per

¹¹ In separate regression results, we find a q_0 value of 36% for movies in the top quartile of popularity, and a corresponding value of 8% for movies in the bottom quartile of popularity.

¹² Estimates for holiday dummies are not shown in Table 8 to save space. The complete estimates are available on request from the authors.

week per movie in the broadcast window as a result of discovery. This value is close to the corresponding value of 464 DVDs per week per movie that we estimated from our reduced form specification (1).

We can now use these parameter estimates to compute the estimated proportion of consumers who have discovered the movies in our sample at the time its broadcast window begins (q_{brd}). We then compare that number to actual DVD sales up to the broadcast window for each movie to obtain counterfactual sales (=actual DVD sales/ q_{brd}) if all potential consumers had discovered the movie at the broadcast window. Table 7 reports these values for movies at different percentiles (based on box-office sales) in our sample.

Table 7: Counterfactual sales under full information

| Percentile | Movie | Observed Sales up to broadcast window | Full information sales | Difference |
|------------|--------------------|---------------------------------------|------------------------|-----------------|
| Max | Kung Fu Panda | 5,572,200 | 5,605,835 | 33,635 (0.6%) |
| 0.90 | Eagle Eye | 1,708,240 | 2,004,977 | 296,737 (17.4%) |
| 0.75 | Tale of Despereaux | 1,074,330 | 1,257,998 | 183,668 (17.1%) |
| 0.50 | Pride and Glory | 406,240 | 520,154 | 113,914 (28.0%) |
| 0.25 | Fired up! | 98,256 | 135,713 | 37,457 (38.1%) |
| 0.10 | Reservation Road | 54,349 | 66,850 | 12,501 (23.0%) |
| Min | How to Rob a Bank | 5,245 | 7,387 | 2,142 (40.8%) |

Table 7 indicates that the movies in the top percentile of box-office sales are not substantially undersold in the DVD window. Almost all consumers are informed about the quality of top decile “hits” before they reach the cable broadcast window: Full information sales for these titles are less than 1% higher than actual sales. However, we find a much larger increase in “full information” DVD sales in all other deciles, suggesting a higher scope of discovery due to broadcast for these movies. We find that at the time of the broadcast window roughly 89% of the consumers are aware about the movies in top quartile but only 77% are aware about the movies in the bottom quartile.

It is also pertinent to note that preferences for movies decline substantially with time. Table 6 reports the average preference decay parameter (δ_t) of -0.132, indicating that preferences for a movie decline by 12% [$1 - \exp(-0.132)$] from week to week. Recall that the DVD sales in a period are a product of two probabilities: the probability of purchase given discovery (preference for movie) and the probability of discovery. So if discovery happens in a later period, sales may increase very little despite higher discovery because of the very low consumer preference for the movie in later periods. In other words, the affect of discovery on movie sales also depends on the time of discovery, making the timing of the broadcast window for movies a very important managerial question.

5. Discussion

While “Long Tail” markets have been observed for some media products such as books, sales of movies remain concentrated in a relatively small number of hits. The literature (Elberse and Oberholzer-Gee 2007, Elberse 2008) provides two main explanations for this concentration in movie sales: (1) heterogeneity in quality and (2) increasing returns from the social nature of movie demand.

In this paper we attempt to develop an additional explanation for the observed skewness in movie sales: incomplete information about movie quality. We test this explanation in the context of the pay-cable release window for movies. The pay-cable window is useful for our analysis for several reasons: First, unlike theater sales, DVD sales, or video-on-demand sales, the pay-cable broadcast channel is the first distribution channel for a movie where there is no “per-item” cost of viewing additional movies, allowing for easier sampling of movies. Second, the timing of the broadcast on the pay-cable channel is such that we can isolate the effect of the broadcast from the effect of other changes in distribution that occur in this window. Finally, the nature of the licensing agreements between pay cable channels and studios reduces selection bias or timing bias that might otherwise exist between popular and less popular movies.

Our analysis shows that movie broadcasts in the pay-cable window significantly increase DVD sales among less popular movies and thereby reduce the skewness of movies sales. To illustrate this change, in our data we see that, prior to broadcast, the top 10% of movies in our sample account for 48% of total sales (the same proportion reported by Elberse 2008) whereas immediately after broadcast the top 10% of movies account for only 35% of total sales.

We argue that consumers might be poorly informed about the true quality of movies because movies are a classic “experience good” that must be consumed to be fully evaluated, and because the nature of movie distribution is such that consumers are likely to be exposed to a relatively small set of available movies. Currently, movies are initially released exclusively through “brick-and-mortar” theaters, as opposed to through “Long Tail” channels. This means that during the theatrical window consumers are only able to view a relatively small number of movies, and as a result studios have incentives to only promote a small number of movies. However, as movies enter the pay-cable window, pay-cable subscribers are able to sample a wider-variety of movies (without incurring an additional “per movie” cost), and our data suggest that this causes their purchase behavior to shift toward more obscure and previously less commercially successful titles.

We then explore this shift in more detail by developing a movie discovery model to precisely estimate the differential discovery of popular and less popular movies. We find that by the time movies reach the broadcast window, 89% of potential consumers have already discovered movies in the upper quartile of

popularity, leaving little scope of its discovery during the broadcast window. In contrast, less popular movies have a much larger scope of discovery during its broadcast and hence we observe higher increases in DVD sales for such movies from discovery in the pay-cable window.

Our research illustrates the importance of product discovery in markets with frequent inflow of new products like movies, music, and books. These results are particularly important as studios begin to experiment with new “digital” distribution channels and alternative distribution windows. For example, MGM, Paramount, and Lionsgate have recently started to put their movies onto a new pay-cable channel called Epix *before they enter the DVD window*. Likewise, Magnolia Pictures released their movie “All Good Things” on cable VOD a month before it was released in theaters and the movie sold very well at both the channels. Our results suggest that these experiments with distribution in digital channels with a capacity to make larger product variety available to wider consumer base (than traditional movie channels such as theaters) may promote movie sales on other contemporary channels. These channels may especially make “Long Tail” movies more commercially viable than they are in the current distribution structure. Our research also has direct implications on movie broadcast windows. In particular, our results suggest that studios would benefit by changing the pay cable window so that movies that were less successful at the box office show on pay cable channels before their more successful movies do..

Despite a rich dataset and robust empirical tests, our study is not without limitations. First, because output deals don’t cover movies from very small studios and documentary movies, our sample is not perfectly representative of all released movies. Second, although we have shown that the entry time in the broadcast window for our sample of movies is not related to the level of box-office sales, the level of DVD sales, or the decay rate of DVD sales, we cannot completely rule out the possibility that pay cable channels are strategically choosing the broadcast timing of movies. Finally, we have also not been able to separately estimate the information spillover effect and the cannibalizing effect of movie broadcast during the broadcast period. In this regard, one interesting extension of our research would be to estimate the blackout effect for cable VOD and electronic channels separately.

References

- Ainslie, Andrew, X. Dre`ze, Fred Zufryden. 2005. Modeling Movie Life Cycles and Market Share. *Marketing Science* 24(3) 508-517.
- Anderson, Chris. 2004. "The Long Tail." *Wired Magazine*, October 2004.
- Anderson, Simon and Regis Renault. 1999. Pricing, Product Variety, and Search Costs: A Bertrand-Chamberlin-Diamond Model. *RAND Journal of Economics* 30(4) 719-735.
- Bakos, Yannis. 1999. Reducing Buyer Search Costs: Implications for Electronic Marketplaces. *Management Science* 43(12) 1676-1692.
- Bailey, J., G. Gao, W. Jank, M. Lin, H. C. Lucas, S. Viswanathan. 2008. The long tail is longer than you think: The surprisingly large extent of online sales by small volume sellers. Working Paper, Smith School of Business, University of Maryland, College Park, Maryland.
- Becker, A. 2007. Universal Pictures and HBO Agree to Long-Term Extension of Film Output Deal. *Broadcasting and Cable*, April 11.
- Brown, Jeffrey R. and Austan Goolsbee. 2002. Does the Internet Make Markets More Competitive? Evidence from the Life Insurance Industry. *Journal of Political Economy* 110(3) 481-507.
- Brynjolfsson, Erik, Yu J. Hu, and Michael D. Smith. 2003 Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety at Online Booksellers. *Management Science* 49(11) 1580-1596.
- Brynjolfsson, Erik, Yu J. Hu, and Michael D. Smith. 2006. From Niches to Riches: The Anatomy of the Long Tail. *Sloan Management Review* Summer 2006, 47(4) 67-71.
- Brynjolfsson, Erik, and Michael D. Smith. 2000. Frictionless Commerce? A Comparison of Internet and Conventional Retailers. *Management Science* 46(4) 563-585.
- Brynjolfsson, Erik, Yu J. Hu, and Duncan Simester. 2010. Goodbye Pareto Principle, Hello Long Tail: The Effect of Search Costs on the Concentration of Product Sales. Working paper, MIT Sloan School of Management, Cambridge, Massachusetts.
- Cabral, L.M.B. 2000. Stretching Firm and Brand Reputation. *Rand Journal of Economics* 31(4) 658-673.
- Cachon, Gerard, Christian Terwiesch, and Yi Xu. 2008. On the Effects of Consumer Search and Firm Entry on Multiproduct Competition. *Marketing Science* 27(4) 461-473.
- Choi, J.P. 1998. Brand Extension as Informational Leverage. *Review of Economic Studies* 65 655-669.
- Elberse Anita, Jehoshua Eliashberg. 2003. Demand and Supply Dynamics for Sequentially Released Products in International Markets: The Case of Motion Pictures. *Marketing Science* 22(3) 329-354.
- Elberse, Anita and Felix Oberholzer-Gee. 2007. Superstars and Underdogs: An Examination of The Long Tail Phenomenon in Video Sales. *Marketing Science Institute* 4: 49-72.

- Elberse, Anita. 2008. Should You Invest in the Long Tail? *Harvard Business Review*, July-August, 1-9.
- Epstein, E.J. 2005. *The Big Picture: The New Logic of Money and Power in Hollywood*, New York: Random House.
- Fleder, D., K. Hosanagar. 2009. Blockbuster culture's next rise and fall: The impact of recommender systems on sales diversity. *Management Science* 55(5) 697–712.
- Frankel, D. 2009. Starz output deals keep films flowing: Pacts with Sony, Disney help fill 16 channels. *Variety*, March 19.
- Goeree, M.S. 2008. Advertising in the U.S. Personal Computer Industry. *Econometrica* 76(5) 1017-1074.
- Hann, Il-horn, Eric Clemons, and Lorin Hitt. 2003. Price Dispersion and Differentiation in Online Travel: An Empirical Investigation. *Management Science* 48(4) 534-549.
- Hedrick, Ken, A. Sorensen 2009. Information and the skewness of Music Sales. *Journal of Political Economy* 117(2) 324-369.
- Krider Robert E., Charles B Weinberg. 1998. Competitive Dynamics and the Introduction of New Products: The Motion Picture Timing Game. *Journal of Marketing Research* 35(1) 1-15.
- Lehmann, Donald R., Charles B Weinberg. 2000. Sales through Sequential Distribution Channels: An Application to Movies and Videos. *Journal of Marketing* 64(July) 18-33.
- Luan, Jackie Y., K. Sudhir, 2006. Optimal Inter-Release Timing for Sequentially Released Products. Working paper. Yale School of Management, Yale University, New Haven, Connecticut. Available at http://faculty.som.yale.edu/ksudhir/workingpapers/SeqRelease_Sep13_KS.pdf (last accessed April 26, 2011.)
- Nelson, Phillip. 1970. Information and Consumer Behavior. *Journal of Political Economy* 78(2) 311-329
- Nelson, Phillip. 1974. Advertising as Information. *Journal of Political Economy* 82(4) 729-754
- Prasad, Ashutosh, Bart Bronnenberg and Vijay Mahajan .2004. Product Entry Timing in Dual Distribution Channels: The Case of the Movie Industry. *Review of Marketing Science* 2(1).
- Sawhney, Mohanbir S., Jehoshua Eliashberg. 1994. Modeling goes to Hollywood: Predicting Individual Differences in Movie Enjoyment. *Management Science* 40(9) 1151-1173.
- Sawhney, Mohanbir S., Jehoshua Eliashberg. 1996. A Parsimonious Model for Forecasting Gross Box-Office Revenue of Motion Pictures. *Marketing Science* 15(2) 113-131.
- Smith, Michael D., Rahul Telang. 2009. Competing with Free: The Impact of Movie Broadcast on DVD Sales and Internet Piracy. *MIS Quarterly* 33(2) 321-338.
- Wooldridge, J. 2002. *Econometric Analysis of Cross-Section and Panel Data*. Cambridge, MA: MIT Press.

Appendices

Appendix-A

To consistently estimate the treatment effect, the entry of movies into the broadcast window must not be systematically related to their commercial success in the box-office or DVD windows. In specification (1) this means that after controlling for time invariant movie characteristics which may affect DVD sales, like genre, movie quality, and the studio promoting the movie, the treatment (i.e. the timing of entry into the broadcast window) must be uncorrelated with commercial success across movies. In Section 3, we argued this is plausible in our setting given the characteristics of output deals between studios and pay-cable channels. In this appendix, we explicitly test whether this is true.

Table A1: Estimates of Cox Proportional Hazard Model for Box-office and DVD sales

| <i>Dependent variable - Weeks between DVD Release and Broadcast</i> | <i>Coefficient estimate</i> | <i>Coefficient estimate</i> | <i>Coefficient estimate</i> |
|---|-----------------------------|-----------------------------|-----------------------------|
| Box-office sales in Millions | -0.001 (0.001) | | -0.001 (0.001) |
| DVD sales in Millions | | 0.017 (0.042) | |
| Movie estimated budget in Millions | | | 0.0002 (0.002) |
| Small studio | -0.289 (0.191) | -0.237 (0.158) | -0.381 (0.231) |
| “Action & Adventure” Genre | -0.175 (0.194) | -0.195 (0.194) | -0.178 (0.192) |
| “Drama” Genre | -0.032 (0.157) | -0.014 (0.156) | -0.089 (0.192) |
| “Comedy” Genre | 0.033 (0.151) | 0.049 (0.151) | -0.023 (0.182) |
| N | 314 | 314 | 227 |

Note - ***, **, * = statistically significant at the 1%, 5% and 10% levels (two-sided test), respectively. Standard errors in parenthesis

To do this, in Table A1, Column 1 we first estimate a Cox Proportional Hazard model¹³ with the number of weeks between the DVD release and the beginning of the broadcast window as the dependent variable, and box-office sales as an independent variable, along with control variables for genre and the type of studio that promotes the movie (“small studio” takes on a value 1 if the movie is from a smaller studio, and 0 if it is from one of the seven major studios: Sony, Warner Brothers, Lionsgate, Fox, Paramount, Disney, and Universal). We then estimate the same model in Column 2, changing only the dependent variable from box-office sales to DVD sales up to the broadcast window. In column 3 besides the box-office sales, we add the estimated movie production budget as a covariate to account for any possibility of high budget but unsuccessful movies entering the broadcast window early¹⁴.

¹³ We get similar insignificant coefficient estimates for box-office sales with Weibull Proportional Hazard model. These results are available from the authors upon request.

¹⁴ We could only get production budget for 227 movies out of our full sample of 314 movies

In Table A1 we find all the coefficients of interest (box-office sales in Column 1, DVD sales in Column 2, and box-office sales and movie budget in column 3) as insignificant. This suggests that, after controlling for time invariant movie characteristics, the entry of movies into the broadcast window is not systematically related to its commercial success or its production budget. In other words, consistent with the available industry information outlined above, a movie’s entry into the broadcast window is not systematically related to box-office or DVD success.

We also test for natural variations in the lag between the blackout period and the broadcast of a movie in the broadcast window across our sample of movies. The concern here is that if pay-cable channels systematically broadcast successful movies in earlier weekends of the month, our blackout and broadcast coefficients would be biased. To test this, we define an indicator variable “*latelag*” which equals 0 if the movie is broadcast in the first two weekends of the month of the broadcast window and 1 otherwise, and estimate the following linear probability model (5):

$$Latelag_i = \beta_0 + \beta_1 \times Boxoffice_i + \beta_2 \times Studio_i + \beta_3 \times Action_i + \beta_4 \times Comedy_i + \beta_5 \times Drama_i + \varepsilon_i \quad (5)$$

Table A2 reports the resulting coefficient estimates. In these results, the coefficient estimate for box office sales is statistically insignificant, suggesting that more successful movies are not shown earlier in a month than other movies are.

Table A2: Estimates of Linear Probability Model for Box-office sales

| <i>Dependent Variable: No. of weeks before entry in broadcast window</i> | <i>Coefficient Estimates (Std. errors in parenthesis)</i> |
|--|---|
| Box-office sales in Millions | -0.001 (0.001) |
| Small studio | -0.195 (0.165) |
| “Action & Adventure” Genre | -0.014 (0.097) |
| “Drama” Genre | -0.038 (0.072) |
| “Comedy” Genre | 0.073 (0.074) |
| Constant | 0.439*** (0.063) |
| N | 314 |

Note - ***, **, * = statistically significant at the 1%, 5% and 10% levels (two-sided test), respectively

Appendix- B

Next, we analyze whether a movie’s DVD sales decay path is correlated with its entry into the broadcast window. Note that regression model (1) only controls for the average rate of decay in DVD sales across all movies in our sample. So, our estimates of the treatment effect will be biased if deviations of a mov-

ie’s DVD sales decay rates from this average rate are systematically related to their time of entry into the broadcast window. In order to test address this issue, we estimate specification (6) using the first difference of the log of DVD sales as the dependent variable:

$$\Delta \text{Log}(S_{it}) = \tilde{\alpha}_i + \tilde{\beta}_{black} \times D_{it}^{black} + \tilde{\beta}_{broad} \times D_{it}^{broad} + \sum_t \tilde{\delta}_t \times D_t + \sum_{wk} \tilde{\delta}_{wk} \times D_{wk} + \tilde{\varepsilon}_{it} \quad (6)$$

where $\Delta \text{Log}(S_{it}) = \text{Log}(S_{it}) - \text{Log}(S_{it-1})$, and where the other variables have the same meaning as in specification (1). This model estimates the impact of the broadcast window on the proportional rate of change in a movie’s DVD sales from week to week. The advantage of this specification is that heterogeneity in sales levels is still accounted for by first differencing, and the fixed effects, $\tilde{\alpha}_i$ control for unobserved heterogeneity in DVD sales decay rates. Taking this heterogeneity from the error term mitigates concerns about potential endogeneity of the treatment with respect to the shape of a movie’s DVD sales decay path. Table B1 reports the coefficient estimates for specification (6).

Table B1: Estimates for first difference form of DVD sales

| <i>Dependent Variable: $\Delta \text{Log}(\text{DVD sales})$</i> | <i>Coefficient Estimates (Std. errors in parenthesis)</i> |
|---|---|
| Blackout period dummy | 0.007** (0.004) |
| Broadcast period dummy | 0.013*** (0.005) |
| N | 17,189 |
| R sq | 0.747 |
| Movie fixed effect | yes |
| Cluster correction at movie level | yes |
| Heteroskedasticity correction | yes |

Note - ***, **, * = statistically significant at the 1%, 5% and 10% levels (two-sided test), respectively

The positive and significant coefficient estimates for the blackout and broadcast effect in Table B1 above indicates that our results are robust to any possible correlation between the DVD sales decay rates of movies with their time of entry in the broadcast window.

One additional concern could be that our results for the broadcast window effect may be merely due to different decay rates for our sample of movies. (For instance, the DVD sales of less popular movies would have flattened towards the weeks inside broadcast window but the DVD sales for the popular movies may still be declining in this period, and this may be the reason for our observed broadcast window effect). We can test for this possibility by only taking the observations for the pre-broadcast window period and randomly assigning a placebo “broadcast” treatment to the movies prior to their actual broadcast

window. We can then use specification (1) to test whether we get any increase in DVD sales due to this artificial treatment infused in our data. Table B2 reports the coefficient estimates.

Table B2 indicates a small and statistically insignificant coefficient estimate for the placebo treatment effect. This shows that the treatment effect that we observe in our data is not merely due to systematically different decay patterns for the movies in our sample.

Table B2: Estimates for placebo treatment effect on DVD sales

| Dependent Variable: Log(DVD sales) | Coefficient Estimates (Std. errors in parenthesis) |
|------------------------------------|---|
| Placebo treatment dummy | -0.005 (0.042) |
| N | 7452 |
| R sq | 0.967 |
| Movie fixed effect | yes |
| Cluster correction at movie level | yes |
| Heteroskedasticity correction | yes |

Note - ***, **, * = statistically significant at the 1%, 5% and 10% levels (two-sided test), respectively

Appendix-C

As noted in Section 4.2, our estimates of the proportional treatment effect may be biased upward because we have large number of less popular movies in our sample (i.e., movies which have low DVD sales at the time of broadcast window). If it were the case that our estimates are driven by low DVD sales for movies prior to the broadcast window, then we should see positive and significant treatment effect estimates for the popular movies that enter the broadcast window late (as these movies will have lower DVD sales at the time of entry in broadcast window).

Table C1: Estimates for popular movies entering late in broadcast window

| <i>Dependent Variable: Log(DVD sales)</i> | <i>Coefficient Estimates (Std. errors in parenthesis)</i> |
|---|---|
| Blackout period dummy | -0.032 (0.056) |
| Broadcast period dummy | -0.105 (0.105) |
| N | 4,098 |
| R sq | 0.863 |
| Movie fixed effect | Yes |
| Cluster correction at movie level | Yes |
| Heteroskedasticity correction | Yes |

Note - ***, **, * = statistically significant at the 1%, 5% and 10% levels (two-sided test), respectively

We can check for this possibility by re-estimating model (1) after restricting our sample to only those movies that are in the top half of popularity and that enter the broadcast window at least 27 weeks after the DVD release (the median value in our sample). Table C1 reports the coefficient estimates for this regression. Note that we find smaller and statistically insignificant coefficient estimates in Table C1 for both the broadcast and blackout period variables. This suggests that our results are not merely driven by smaller DVD sales for large number of less popular movies in our sample.

We estimated the proportional treatment effect because of highly skewed DVD sales in our sample of movies. This variation in DVD sales may be smaller for movies in a popularity-based quartile in our sample. We therefore estimate the linear average treatment effect separately for the movies in each quartile of popularity in our sample. For this, we use the same specification (1) but with weekly DVD sales as the dependant variable instead of Log of weekly DVD sales. Table C2 reports the coefficient estimates.

Table C2: Estimates for simple treatment effect for least popular movies

| <i>Dependent Variable: DVD sales (Std. errors in parenthesis)</i> | <i>Coefficient Estimates</i> | | | |
|---|------------------------------|--------------------------|--------------------------|---------------------|
| | 1 st Quartile | 2 nd Quartile | 3 rd Quartile | 4th Quartile |
| Blackout period dummy | -4064.39** (1906.26) | 1286.38*** (430.97) | -132.16 (318.75) | 1.36 (43.91) |
| Broadcast period dummy | -673.44 (1704.14) | 1578.85*** (367.23) | 619.56*** (279.18) | 69.390** (30.52) |
| N | 4105 | 4061 | 4221 | 4813 |
| R sq | 0.52 | 0.80 | 0.63 | 0.56 |
| Movie fixed effect | Yes | Yes | Yes | Yes |
| Cluster correction at movie level | Yes | Yes | Yes | Yes |
| Heteroskedasticity correction | Yes | Yes | Yes | Yes |

Note - ***, **, * = statistically significant at the 1%, 5% and 10% levels (two-sided test), respectively

We find a positive and significant coefficient estimate for the linear treatment effect of broadcast for the movies in bottom three quartiles but a small and insignificant estimate for the movies in top quartile of popularity. In all, the results of the linear treatment effect are similar to the proportional treatment effect. This further indicates that our results are not merely due to using a proportional treatment effect.