MODELS OF COMPETITION IN THE U. S. MOTION PICTURE INDUSTRY

A Dissertation
Presented to the Faculty of the Graduate School
of Cornell University
In Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy

by
Anirban Mukherjee
May 2009

MODELS OF COMPETITION IN THE U. S. MOTION PICTURE INDUSTRY

Anirban Mukherjee, Ph. D.

Cornell University 2009

My dissertation builds and estimates models of competition for the U.S. motion picture industry. The motion picture industry provides a rich setting for exploring the effect of competition. First, there is a rapid refresh cycle with new titles being the mainstay of industry activity. Modeling new releases adds to estimation complexity due to more stringent data demands. Second, the industry sees seasonal changes in demand and competition. The frequent entry of new products leads to an emphasis on release timing and pricing. Third, the presence of multiple generations of technological formats leads to time varying cannibalization and substitution across channels. While I study the motion pictures industry, insights generated and methods developed are generalizable to other industries with multiple-channel distribution, critical entry and pricing decisions, and technology platform transitions.

Specifically, I model three elements of competition in this industry. In essay 1, I model how movies compete with other movies in the same channel and other channels (e.g. theaters, rentals, purchases, etc.). I allow for movies within any channel to be substitutes or complements for movies in the same and other channels. In essay 2, I model how release date and price matter for movie competition; the application is to DVD sales channel. In essay 3, I model how studios compete in format wars. The application is to studios planning whether to release movies in DVD and/or VHS and

at what price, in the days when the DVD format was growing, and studios had the choice of waiting for the rivals to do the work of subsidizing the new format.

My dissertation delivers insights on managing marketing mix variables across formats, accounting for seasonal demand, competition and cost differences. The models both allow inference of payoffs, and provide counterfactuals for future actions of managers, accounting for competition. The model frame works described provide methodological advances for model specification and estimation in other industries allowing generalizability. In particular, I contribute to the literature in dynamic games by introducing partial and complete information estimators for games with multiple potential equilibriums played in the data.

BIOGRAPHICAL SKETCH

Anirban Mukherjee was born in Mumbai, India. He moved to Delhi in 1986 and went to the Delhi Public School for grades 1 through 5. He went to Welham Boys'School, Dehra Dun for a year and completed his education at The Doon School, Dehra Dun (353 KA, 1999).

Anirban joined Cornell University in 1999. He was awarded his B. Sc. (cum laude) in 2003, majoring in Electrical and Computer Engineering. He started his doctoral studies in fall 2004 and was awarded a M. Sc. in management in spring 2008. Through the Beautiful Ithaca Gorges, Rivers Endearing Depart: after a decade spent in Ithaca, he looks forward to the tropical climate of Singapore, where he will be starting as an assistant professor in marketing at the Singapore Management University.

To my parents

ACKNOWLEDGMENTS

I thank the Johnson faculty for its commitment to its students. In my five years at Johnson, I never felt unwelcome to say my piece, or at a loss of people or resources to turn to for knowledge or inspiration. My particular thanks to the marketing department for their very generous intellectual and financial support.

My friends have been a source of joy through my life. A special thanks to Lona Fowdur, Lynn Li, Abhijit Patnaik, Bernard Tarr, Joanna Upton and Kevyn Yong, for trysts endured.

I am grateful to my aunt, Sabita Saha, for her kindness and guidance in my early youth; her love remains a bellweather in turmoil.

I thank my adviser, Vrinda Kadiyali, for her guidance. Her support was unwavering in my idiosyncratic journey through the doctorate. She was my rock: emotionally and cerebrally. She has a fiercely independent perspective which carved the bedrock of my dissertation, and I promise to work to build upon the knowledge she imparted.

Last, my unreserved gratitude and love to my parents; I dedicate this dissertation to them. Thank you!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

FORECASTING IN RAPIDLY CHANGING ENVIRONMENTS: AN APPLICATION TO THE U.S. MOTION PICTURE INDUSTRY[1]

1. Introduction

Industry watchers and researchers have discussed how a growing number of industries are facing rapidly changing environments. Example drivers of change are technology, globalization, capital market pressures, and new ownership structures via mergers and spin-offs. In changing environments, long histories for a product are unavailable or not reliable in changing environments, managers in such industries face difficulty in forecasting what lies ahead. Our goal in this paper is to provide a tool for sales forecast in rapidly changing environments using relatively short histories.

The application is to the U.S. motion picture industry.[2] The high costs of movie production and frequent failures at the box office raise the stake in forecasting movie

---

[1] This essay is co-authored with Vrinda Kadiyali, a Professor of Marketing and Economics at the S. C. Johnson Graduate School of Management, Cornell University, 385 Sage Hall, Ithaca, NY 14853. We thank seminar participants at Cornell University for comments. We also thank Nielsen EDI, Paul Kagan and Associates, Nielsen VideoScan, and Home Media Retailing for providing data for this study.

[2] While our application here is to movies, our model is equally applicable to other products with similar characteristics. Examples include television shows (on television, off-network syndicates, and home videos), music (audio and video singles in hard copy, on-line, as part of an album and then in broader compilations), and even books (hard bound and then in paperback and e-books.). The fashion industry also has similar characteristics of short lifecycles, seasonality, and cross-channel competition.

revenues. Several factors make forecasting movie sales inherently difficult: the difficulty in quantifying what makes a movie successful (the role of unobservables in the data), short lifecycles of the product, and facing different competitive sets each week. Changing consumption patterns, driven by improvements in technology, reduce the usefulness of historical data in forecasting.

We build a model that forecasts revenue for movies before they are launched and in four separate channels: theatrical release, DVD sales, VHS sales and rentals, by title, by week. Our model extends the Multiplicative Competitive Interaction model of Cooper and Nakanishi (1988) to account for unique features of this market. Specifically, we account for (1) observed and unobserved movie characteristics, (2) seasonality of demand, (3) competition within and across multiple distribution channels, (4) market expansion, substitution and/or complementarity between movies inside and across distribution channels. This involves significant data efforts. This also involves overcoming an important methodological issue. We allow for general correlation in error structures to do (4). Therefore, we cannot use movie characteristics in one channel as instrumental variables for movie performance in another channel because characteristics of a movie in one channel are likely to be systematically correlated with error terms in the other channel. Therefore, we modify existing methodology (Chiou, 2007) to estimate our model. We find that our model improves forecasts in the sequential distribution channels, with lower mean squared error in out of sample validation than extant models.

The rest of the paper proceeds as follows. In the next section we discuss issues in forecasting revenue in the movie industry. Next, we provide a literature overview. In section 4 we discuss our model in greater detail. In section 5, we provide details on the empirical application. The last section concludes.

2. Issues in Forecasting Movie Revenue

As mentioned in the introduction, movies are highly differentiated products, and it is nearly impossible to gather enough data to explain all the determinants of revenue success. For example, genre is likely to be a useful (observed) characteristic of a movie in any forecasting exercise. However, a movie's plot likely also matters for revenue forecasts. However, data on plots is not easy to quantify as an explanatory variable. We want to model competition among movies within and across channels. This raises additional issues related to unobservables. Some movies may be complementary; the release of a popular movie on DVD may increase visits by to retailers, and hence increase sales of other titles. Promotions for titles with similar plots might have positive (or negative) spillover effects across titles. In our model, we control for unobservables and in modeling competition across channels, allow for market expansion (or contraction), complementarity and substitution within and across channels.

Consider next the issue of short-life cycle of a movie in theatrical release, and the large number of alternatives available in theatrical and across other channels. This leads to intense market share competition (Epstein, 2005). Consumers who do not watch a title in a channel, may either substitute the title entirely and never watch the title, or watch the title in a different channel, substituting across channels. The short life-cycle in the primary channels increases the need for accurate forecasts of revenue for the first weeks post release. Should forecasts of demand be inaccurate, there is not much time to improve non optimal promotion and distribution strategies.

Further, the large proportion of sales in the first weeks post-release (Ainslie, Dreze and Zufryden, 2007), causes data constraints in forecasting. In more prevalent approaches in marketing focused on established brands and products, long data histories are used to estimate brand-specific parameters when forecasting future demand. A managerially relevant forecasting model for the movie industry must only use the set of observables known to a manager at the time of forecasting and cannot use prior weeks' revenues to forecast later week revenues. Thus, our goal is to forecast sales by title using data available prior to release of the title in the channel.

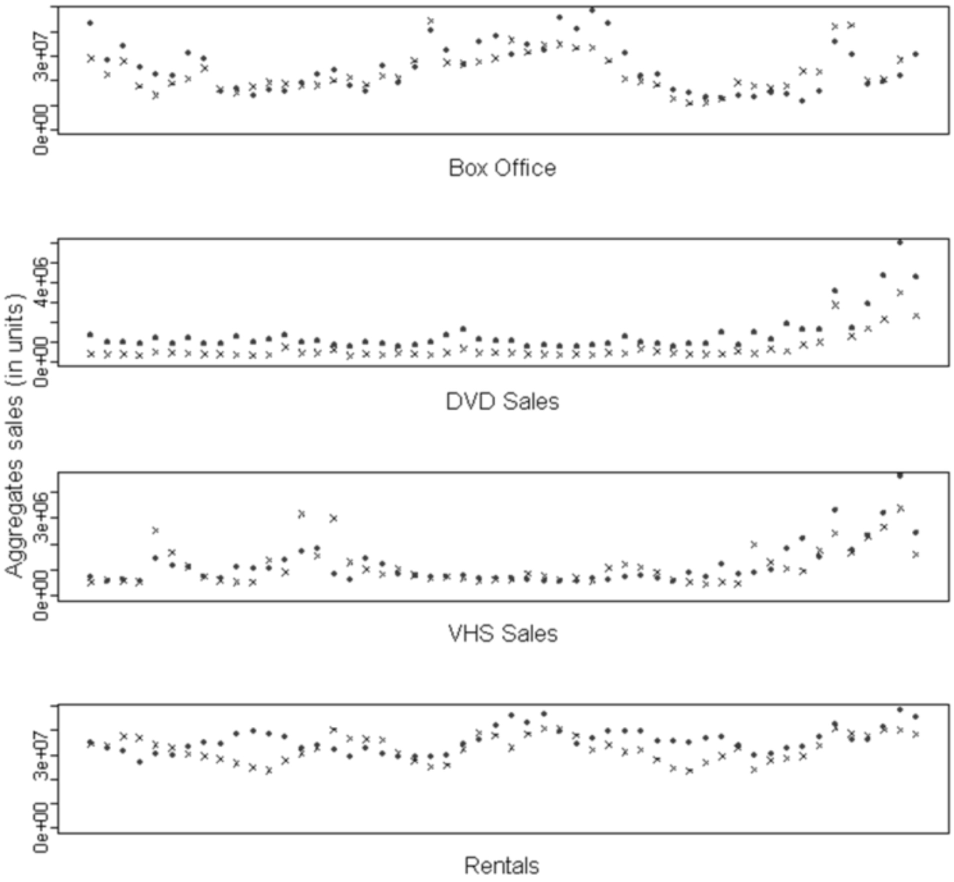

Figure 1: Aggregate Weekly Revenue

Another issue in movie forecasting is the seasonality of demand for movies in both primary and secondary channels (see Figure 1). Studios account for this seasonality in their theatrical and secondary channel distribution (see Einav 2003), with the most highly awaited movies released in weeks of peak demand. For accurate forecasts, it is important to account separately for changing seasonal demand and market expansion due to new releases. That is, are more movies released in a high-demand season because of the high demand, and/or do they cause demand to increase given the greater variety (and possibly quality) of movies available to viewers in these weeks? Our model separately controls for both effects in each distribution channel.

Over the last 2 decades, the movie industry has faced a rapidly changing environment. The advent of the video cassette (particularly VHS) in 1984 and then the Digital Versatile Disc (DVD) in 1997, made home-viewing possible. These technologies led to new channel partners (and new complementors via merchandizing deals), altering competitive landscapes. Secondary channels and merchandizing deals grew in importance to the industry.[3] Expectedly, the entry of new players in the industry landscape, led to changes in pricing and distribution policies in channels. For example, in the period of 1996-1999, the introduction of a new rental revenue sharing mechanism reduced inventory risk allowing rental chains to keep more copies of a title in stock (see Mortimer, 2004). Other changes include the growth of the online rental

---

[3] In 1980, the industry made approximately 30% of its revenues from the domestic box office, and 7% of its revenues from home video (including rentals and sales); in 2000 the industry made approximately 15% of its revenue from box office and 38% of its revenue from home video (Vogel, 2004).

and sales channels, and peer to peer movie piracy (Smith and Telang, 2006).

We build a model taking in to account the issues mentioned above. Empirically, the time period for which we forecast the secondary channel revenues are the last 4 months of 2001. Rapid change in an industry environment implies that the window of past data useful in forecasting is likely to be small. Cognizant of the data limitations imposed by the changing environment in the movie industry, we estimate and benchmark models using the data from January 2000 to the last week of June 2001.

3. Literature Review

The movie industry has invited considerable attention from several marketing scholars (see Eliashberg, Elberse and Leenders, 2006, for a summary). We organize our review below as models of single movie performance, models with competition within a channel, and models with competition across channels. Our model is in the final category.

Consider papers that examine single movie performance with no competition within and across channels. Sawhney and Eliashberg (1996) propose a model (BOXMOD) for box office performance. Eliashberg, Jonker, Sawhney, and Wierenga (2000) propose a model (MOVIEMOD) for predicting pre release awareness, adoption intent and cumulative penetration in consumers. Neelamegham and Chintagunta(1999) and

Elberse and Eliashberg (2003) focus on international box office receipts.[4] These papers model sales as a function of observable characteristics (such as budget) and past performance of a movie. Our formulation for market attractiveness is informed by these papers, and we compare our model performance to BOXMOD, modified for the sequential channels.

Other single-movie papers have modeled other important industry features. Shugan (1998) looks at the impact of the production team on box office success. Krider and Weinberg (1998) discuss competition when faced with seasonal demand variations. Radas and Shugan (1998) outline an approach for including seasonal trends in estimating demand curves by taking a transformation of time. Important insights from this literature for our model are the decline of receipts over weeks post release, and the impact of print and advertising expenditure on box office performance.

Consider next papers that model model within-channel competition. Three approaches have been used here. Swami, Eliashberg, and Weinberg (1999) study multiplex screen allocation decisions and formulate a model to optimize exhibitor scheduling. Ainslie, Dreze and Zufryden (2007), hereafter ADZ, build on the BOXMOD model and study the lifecycle of a movie at the box office, measuring the substitution effects of competition within a channel. Einav (2007) presents an empirical analysis of release timings in the U.S. movie industry, studying both seasonality and competition. In a

---

[4] Lee, Boatwright and Kamakura (2003) draw upon single movie prediction models and specify a hierarchical Bayesian model, to forecast sales of music albums, prior to their launch. Whilst the industry of their study is different, the research question and modeling challenges faced are similar.

companion paper, these estimates are used to study the timing game, and optimal timings calculated for the industry (Einav, 2003).

Our model is similar to ADZ and Einav (2007) in allowing for flexible competitive structures and seasonality. Expanding on these models, we include market expansion and allow for more flexible revenue patterns beyond the two-way classification of steady decay blockbuster movies and sleeper movies. It is not clear ex-ante if only two types of patterns are present in secondary channels, and how the primary channel revenue patterns might change when we consider the substitution/complementarity of secondary channels and the longer availability of a title in a channel. Therefore, our model can be seen as a general case of their model. Also, unlike these papers, we cannot use long datasets to estimate relatively stable traits like seasonality, and our model of competition within and across channels poses additional data gathering and methodological tasks.

A third relevant stream of literature examines competition across channels. Here, researchers have examined competition for any given movie across channels, without modeling competition amongst movies within a channel. An example is Lehmann and Weinberg (2000), who develop a model of the optimal time to enter a second channel for any movie. They calculate optimal release timings in rentals accounting for the cannibalization of sales from theatrical release. They do not study cannibalization across the secondary channels. Prasad, Bronnenberg, and Mahajan (2004) use an analytical model to study the effect of consumer expectations on optimality of the timing decision. The duration between releases is treated as an unwritten covenant of the industry, shaping customer expectations. In contrast, the optimal timing policy for a distributor depends on current expectations, leading to an impetus to cheat and release early. Luan and Sudhir (2007) model the impact of cannibalization of sales and rentals of movies, on box office revenues, accounting for forward looking behavior of the consumer at the theatre.

Two papers study movies in a multiple channel setting. First, Hennig-Thurau et al, 2007, use individual level discrete choice data to study the effect on studio profitability of different configurations of sequential distributional channels, optimizing release timings across these channels. As their goal is to study hypothetical configurations vastly different from current market conditions, they use conjoint data to model channel substitution, without accounting separately for either complementarities or market expansion. Second, Chiou (2007) models seasonal demand variation in secondary channels, controlling for competitive interactions within the rental revenue channel and in DVD and VHS sales. While Chiou's model is the closest to ours, we cannot use her estimation methods in a multi-channel setting where unobservables across channels might be correlated. Therefore, we develop methodology more appropriate to our setting. More on that in section 4 below.

4. Revenue Prediction Model

We first outline the market attraction function. Second, we describe our model of market share across channels. We conclude by describing the estimation methodology needed to forecast revenues.

4.1: Market Attraction

In the MCI model, the market share of a product is a function of the ratio of the market attraction of the product to the sum of the market attractions of all products. We define our market attraction a for movie m, in channel k, in week w, year y as below:

$$\ln(a_{mkwy}) = \ln(\delta_{mkwy}) - \tau_{kw} + \xi_{mkwy}$$

The components of the attraction are as follows. The deterministic component (δ_{mkwy}) is a function of observed variables. The specification of this function and a description of its components in our empirical application, are provided in section 5.2. τ_{kw} is a weekly unobservable shock common to all products in a given channel and week. It is likely that the opportunity cost of watching a movie in the theatre, and/or on video rental, varies by week. For example, opportunity cost is low during holidays and higher during working days. In the summer, the costs of driving to a store can be considerably different from that in the winter. Unobserved product attributes are modeled as product specific shocks ξ_{mkwy}. These include movie plot, and the psychological and informational setting of a consumer (Eliashberg and Sawhney, 1994, and Neelamegham and Jain, 1999).

We do not specify the distribution of unobserved characteristics, ξ_{mkwy}. Instead we only restrict its first moments in a quasi likelihood specification for identification, setting $E\left[\xi_l \in \Xi_{wy} \mid x_{mkwy}, \tau_{kw}\right] = 0, \forall \xi_l \in \Xi_{wy}$, with Ξ_{wy} a vector of all product shocks is in week w, year y. Our specification is flexible enough to accommodate three important sources of covariance in the movie industry. First, we expect contemporaneous correlation between movie specific shocks in a given channel, in each week. For instance, movies released in summer might share similar characteristics. Second, shocks of movies of movies released in different weeks may show different contemporaneous correlation. For instance, while the older summer blockbuster titles in September would continue to exhibit summer-movie correlation, newer movies released in September might have fall-movie characteristics. Hence the unobservables of new releases will have a different correlation from older releases. Third, noting the rapid decrease in market attractiveness post release, and the simultaneous release of a title in DVD sales, VHS sales, and rentals, shocks from a title will exhibit serial correlations across different weeks of demand, and across different channels.

Our quasi likelihood specification allows for these possibilities. Compare it to a (hierarchical) Bayesian specification, similar to Lee, Boatwright and Kamakura (2003). The latter requires explicitly modeling the covariance matrix of the unobservables. The likelihood of the observed data, a key component of such a model, cannot be formed unless one specifies the relationship between observations in different channels and across different weeks. Without significantly restricting degrees of freedom for the covariance matrix, the limited number of observations on each movie makes for imprecise estimates of this matrix. Instead, we choose to use a more flexible specification that provides less efficient but consistent estimators.

4.2: Market Share Model

Let $\mathbb{C}_{kwy}$ be the choice set of movies in channel k, week w, and year y. In MCI, the market share of movie m (denoted ms_{mkwy}) is written as

$$ms_{mkwy} = \frac{\exp\left(\ln(a_{mkwy})\right)}{\sum_{i \in \mathbb{C}_{wy}} \exp\left(\ln(a_{ikwy})\right)}$$

Our model applies to any nesting structure with as many channels; in our empirical application we model four channels for which we have data. For each additional channel, the number of coefficients grows linearly. We write the in-channel market share, percentage of cumulative sales of all movies in the channel of movie m in channel k, week w, year y (denoted $ms_{mwy|kwy}$) as

$$ms_{mwy|kwy} = \frac{\exp\left(\frac{\ln(a_{mkwy})}{(1-\rho_k)}\right)}{D_{kwy}}$$

$$where\ D_{kwy} = \sum_{i \in \mathbb{C}_{kwy}} \exp\left(\frac{\ln(a_{ikwy})}{(1-\rho_k)}\right)$$

We allow the attraction functions of new title releases to be correlated. Thus we account for complementarities between titles, for instance if such titles were sequels, and crowding out/ negative externalities exerted by simultaneous releases. We constrain ρ_k to be strictly between 0 and 2 to maintain consistency with the literature in discrete choice models. The cumulative market share of all titles in a channel k is:

$$ms_{kwy} = \frac{D_{kwy}^{1-\rho_k}}{1+\sum_{j \in \{Channels\}} D_{jwy}^{1-\rho_j}} = \frac{D_{kwy}^{1-\rho_k}}{D_{wy}} \tag{1}$$

ρ_k is a channel nesting parameter that controls market expansion, cannibalization and substitution. The release of a new title in channel k increases D_{kwy}, with (1) determining the new total channel revenue after market expansion. The market attraction function is scaled by the channel nesting parameter in the market share model, thereby controlling substitution of movies within a channel. Between the values of 0 and 1, the derivative of the market share function with respect to the nesting parameter is positive, suggesting an increased sensitivity to differences in the attraction functions. Between the values of 1 and 2, the derivative of the market share function is negative; suggesting that the market share function reverses direction. Cannibalization is controlled by the nesting parameters of the channels that the title has been released in. Release of a new title into channel k increases D_{kwy} with (1) determining the new total channel revenue for channels in which the title was currently available, after cannibalization.

4.3: Deriving the Revenue Forecasting Equation

To forecast revenue we have to account for both τ_{kw} and ξ_{mkwy}. We develop a novel forecasting model in two steps. First we integrate product shocks, ξ_{mkwy}. Second, we substitute for channel specific, unobserved time shocks, τ_{kw}.

Define Eq_{wy} as the expected aggregate volume of titles bought or rented (across all channels) in week w, year y; Eq_{kwy} as the expected aggregate volume of titles bought or rented (channel specific) in channel k, week w, year y; Eq_{mkwy} as the expected volume of movie m, bought or rented in channel k, week w, year y.

We predict revenue by setting $Eq_{mkwy} = Ems_{mkwy}M$. While M is set to a large number and never observed, ms_{mkwy} is a function of stochastic variables $\left\{\xi_{mkwy}\right\}$. Note that $ms_{mkwy}\left(\xi_{mkwy}\right)$ is strictly monotonically increasing in ξ_{mkwy}, continuous and differentiable everywhere. The first order Taylor expansion around $E\left[\xi_l \in \Xi_{wy}\right] = 0, \forall \xi_l \in \Xi_{wy}$ leads to $E\left[ms_{mkwy}(\xi_i \in \Xi_{wy})\right] \approx ms_{mkwy}(E\left[\xi_i\right] \in \Xi_{wy})$. Hence, we substitute the expected unobserved product shock for the unobserved product shock in the first step. In the second step we substitute for the unobserved time shock. Index channels as B for box office, D for DVD sales, V for VHS sales, and R for rentals. Define:

$$D_{kwy} \triangleq \sum_{i \in C_{kwy}} \exp\left(\frac{\ln(\delta_{ikwy}) - \ln(\tau_{kw})}{1-\rho_k}\right)$$

$$\kappa_{kwy} \triangleq \sum_{i \in C_{kwy}} \exp\left(\frac{\ln(\delta_{ikwy})}{1-\rho_k}\right)$$

$$\gamma_{wy} \triangleq \frac{(\kappa_{Bwy})^{1-\rho_B}\tau_{Bwy} + (\kappa_{Dwy})^{1-\rho_D}\tau_{Dwy} + (\kappa_{Vwy})^{1-\rho_V}\tau_{Vwy}}{\tau_{Rwy}} + (\kappa_{Rwy})^{1-\rho_R}$$

Thus we write

$$Eq_{wy} = \frac{\gamma_{wy}}{\tau_{Rwy} + \gamma_{wy}} M \tag{2}$$

The goal of this paper is to derive a model for predicting the sales of a title, in a given week, in a given channel, before release of that title in that channel. The above expressions derive total sales in a week, summed over all channels, as a function of the total market size, M, and unobserved parameter τ_{Rwy}.

First, noting that this parameter, τ_{Rwy}, is the same over a particular week in the two year sample, we substitute for the unobserved terms in the equation for the second year, to get:

$$Eq_{w2} = \frac{\gamma_{w2}}{\left(\frac{M\gamma_{w1}}{Eq_{w1}} - \gamma_1\right) + \gamma_2} M \tag{3}$$

Next, we re-arrange and derive an equation for the total sales in a week, in a given channel. From (2) and (3), we write:

$$\begin{aligned} Eq_{Rw2} &= \frac{(D_{Rw2})^{1-\rho_R} Eq_{w2}}{(D_{Bw2})^{1-\rho_B} + (D_{Dw2})^{1-\rho_D} + (D_{Vw2})^{1-\rho_V} + (D_{Rw2})^{1-\rho_R}} \\ &= \frac{(\kappa_{Rw2})^{1-\rho_R}}{\gamma_{w2}} Eq_{w2} \end{aligned} \tag{4}$$

Next, we derive the final equation for sales of a title, in a given week, in a given channel. Without loss of generality, consider forecasts for movie m, in rentals channel subscripted by R, in week w, year 2. Using (4), we get:

$$Eq_{mRw2} = \exp\left(\frac{\ln(\delta_{mRw2})}{(1-\rho_R)}\right) \frac{M(\kappa_{Rw2})^{-\rho_R} Eq_{w1}}{\left(M\gamma_{w1} - \gamma_1 Eq_{w1}\right) + Eq_{w1}\gamma_2}$$

To predict revenue we replace expected sales in year 1 with observed sales in year 1.

$$q_{mRw2} \approx \exp\left(\frac{\ln(\delta_{mRw2})}{(1-\rho_R)}\right)\frac{M(\kappa_{Rw2})^{-\rho_R} q_{w1}}{\left(M\gamma_{w1}-\gamma_1 q_{w1}\right)+q_{w1}\gamma_2} \tag{5}$$

Our final forecasting equation (5) can be used for long term forecasts, predicting revenue on a set of movie characteristics and observables available months earlier, prior to the release of the movie in that particular channel. Note that when using a longer dataset, (5) can be used for each preceding year and an estimate formed from the (weighted) average.

4.4: Two-Step Estimation for Endogenous Choice Sets

Our nested MCI model is similar to a formulation by Chiou (2007) where (1) leads to:

$$\ln(ms_{mkwy}/ms_{0wy}) = \ln(a_{mkwy}) - \rho_k \ln(ms_{mwy|kwy})$$

The last term, in-channel market share, $\ln(ms_{mwy|kwy})$, is endogenous and correlated with the product shock in the attraction function. Commonly, attributes of other titles in the channel that affect the in-channel share and are not correlated with the market attraction, are used as instruments. For instance, Chiou (2007) uses the sum of the characteristics of other products, as instruments.

This estimation strategy requires the attributes of other titles in the channel to not be correlated with the product attraction shock. As discussed earlier, studios time the release of the best movies to be in periods of highest demand. The number of movies released in a given week as well as the cumulative budgets of all movies in the channel in a given week, show strong seasonal patterns. Hence, attributes of films released in the same week are strongly correlated. For instance, blockbusters in theatrical, sales and rentals, are all simultaneously released in the same weeks in

December. Therefore, characteristics of movies in one channel cannot be used as instruments for a market-share model for another channel, as the unobservables might be correlated across channels (e.g. summer-themed movies are released in theatrical and sales channel in summer and winter-holiday themed movies are released in theatrical and sales channels around the winter holidays).

Thus, the release timing game implies that the assumption of exogenously determined choice sets is likely to be inaccurate, biasing the described instrumental variable estimator. For instance, movies released in a peak summer week have systematically larger budgets, and systematically larger average product attractions as they were picked by studios to be summer releases. Conversely, movies with smaller product attraction shocks, should on average, be released in weeks with less competition. Instrumental variable estimates of the nesting coefficient are biased in both cases.

In addition, channel coefficients are identified through the variance across the characteristics over the years, for the same week. Due to the release timing game and the underlying stability of seasonal patterns, such variance remains limited when compared to variance across weeks in a year. For instance, holiday weeks across years have similar cumulative budgets over all releases in a week, as all big budget movies of the year are released in this period. These concerns are amplified when using a shorter dataset, as in our problem, where there are far fewer overlapping weeks over which one can identify the channel coefficients.

Given the issue with instrumental variable, we use a two-step estimation process. In the first stage, we identify the market attractiveness function to scale by regressing a function of market shares on characteristics. In the second stage, we estimate channel nesting parameters by minimizing an objective function formed through prediction errors of the forecasting equation.

Let $\tilde{s}_{jkt}$ be the market share of movie j in channel k, in week w, year y, in quantities. Define the geometric mean of in group market shares as $\ln(s^{g}_{kwy}) = \frac{1}{N_{kwy}} \sum_{i \in C_{kwy}} \ln(\tilde{s}_{ikwy})$ Then from(2):

$$\ln(\tilde{s}_{jkwy}) - \ln(s^{g}_{kwy}) = \ln(\delta_{jkwy}) - \frac{1}{N_{kwy}} \sum_{i \in C_{kwy}} \ln(\delta_{ikwy}) + \xi_{jwy} - \frac{1}{N_{kwy}} \sum_{i \in C_{kwy}} \xi_{ikwy} \quad (6)$$

In the first stage, coefficients from (6) are estimated using Ordinary Least Squares[5] and used in the second stage to find the channel nesting parameters. Conditional on a guess of channel coefficients, we calculate differences in the response to mean price and the seasonal change in demand. Using(5), we can predict the total revenue of a channel in the future weeks.

Having estimated the first stage on the first 84 weeks, we use the next 4 weeks to find channel nesting parameters that minimize the sum of squared errors in the prediction sample.[6] As our objective function does not have analytical derivatives, we utilize a Nelder Mead simplex search followed by the Broyden-Fletcher-Goldfarb-Shannon method (with numerical derivatives), to find the minimum. While we do not prove the

---

[5] Clustering errors by week and using White's correction for heteroskedasticity does not improve fits and/or predictions.

[6] The sum of squared errors and the sum of absolute errors led to similar estimates and predictions.

existence of a unique global minimum, varying starting values we find the algorithm converges to a unique parameter vector. Robustness tests indicate that our prediction results are not affected by the size of the market, as long as we choose M larger than the maximum total quantity of entertainment products sold.

Thus, we develop a new estimation algorithm to account for systematic correlation between weeks of peak demand and the release schedule of better performing movies, and face the burden of estimating without movie-specific parameters. Note that our model of competition is reduced-form. A full structural model of competition that accounts for substitution, complementarity and market expansion is beyond the scope of our research question. Instead, we build a model that has a flexible competitive structure and market attraction formulation, without attempting inference on competitive structures.

5. Empirical Application: Forecasting Weekly Movie Revenues

We describe the data used in our empirical application, forecasting movie revenues in the U.S. motion picture, the operationalization of the forecasting model, and present our results.

5.1: Data

We use data from three distribution channels – the primary channel i.e. theatrical release, and two secondary channels, rentals and sales. As mentioned previously, the data are for January 2000 through December 2001.

We obtained data from Nielsen EDI for nationally aggregate theatrical revenues (and distributional reach) in the first ten weeks of theatrical release for all movies released on box office. In the dataset there are about 40 movies being exhibited across theaters nationwide in any given week. Nielsen Videoscan collects DVD and VHS sales data from retailers at the point of sale. We use data (including weekly sales and price) for the top 500 selling movie titles in each channel, which covers all movies that sell over 300 copies in a format, in a particular week nationally. Other researchers have used this dataset to study movie VHS and DVD sales (Elberse and Oberholzer-Gee, 2006). There are significant differences in lifecycle, pricing and other competitive issues. Therefore, we model the two formats separately. The data does not include Walmart. In our period of interest, Walmart was a major retailer of DVD and VHS that carried a smaller inventory of possible titles than comparable national retailers. Hence, our sample may understate the importance of larger titles and overstate the importance of smaller titles.

Rental data comes from Video Store Magazine's Rental Charts. This source tracks weekly national revenue by title for the top 50 selling titles that week. Video Store Magazine constructs estimates from a panel of suppliers and retailers. Therefore, compared to the sales data from VideoScan, these data might be more inaccurate but relatively unbiased if the panel of suppliers and retailers is representative. Also, unlike the sales data, rental data is not divided by DVD/VHS format.

Pricing policies differ across channels. In movie theatres and home video rentals, prices are almost always uniform (Einav and Orbach, 2007), and revenue shared between the exhibitor/rentailer and distributor We assume a mean box office ticket price of $6.00 per title, and a mean rental price of $2.50 per title for the duration of the study (Hettrick 2000, Vogel 2004) and find our results robust to a range of price means.[7] In DVD and VHS sales, we observe the weighted average price for a title (reported by week, by title) sold in all Discount Mass, Drug & Grocery stores, which account for about 43% of all units sold in the dataset, and assume it to be equal to mean price of the movie across all reporting retailers.[8]

To enrich the forecasting model, we gather data on additional variables by title. We obtain data on print and ad spending (P & A) for each movie at the box office stage from Paul Kagan and Associates. We also use user ratings from Internet Movie Database (http://www.imdb.com) to proxy for user reported quality. We complement user ratings with a summary measure of critics' ratings. Available at Rotten Tomatoes (http://www.rottentomatoes.com), the Tomatometer, captures the percentage of positive critics' reviews for a title. Additionally, sales of a movie may be influenced by the "star" power of the actors and directors involved in the movie (see Elberse, 2006 for a summary of studies on star power). Ulmer (2000) published a list of the top 200 actors and top 10 directors in Hollywood, measured on their "bankability" at the

---

[7] Two major retail pricing schemes existed in DVD and VHS. First, in "sell-through" pricing, retail prices were set with the expectations of direct purchase by consumers and rental stores, and the title distributed widely to media retailers. Second, in "rental-window" pricing, a movie was priced above the average retail price of other VHS or DVD available at retailers, and distribution curtailed to rental stores. While DVDs were introduced in 1997 with "sell-through" pricing, "rental-window" pricing was the norm for VHS cassettes through the late nineties. Further information on theatrical contracts can be found in Einav (2007), and on rental contracts can be found in Mortimer (2004) and Chiou (2007).

[8] A number of retailers, including Target and Kmart but excluding Walmart, are included in the category.

box office, through an industry wide survey of Hollywood professionals. We gather data on all actors and directors featured in a movie, and check if those actors were included in the top 200 and top 10 lists respectively.

Table 1: Summary statistics by channel[9]

	Box Office	***DVD Sales***	***VHS Sales***	***Rentals***
Channel Gross (in $)	3,278,381 (8,270,558)	97,127 (23,577)	82,045 (47,888)	2,030,033 (2,495,021)
Box Office (in $) (for sequential channels)	n/a	80,941,000 (74,159,950)	91,119,160 (75,999,900)	46,447,260 (54,094,390)
Screen-Weeks (for sequential channels)	n/a	14,490 (6,725)	15,666 (6,548)	10,503 (7,260)
Budget (in $)	26,554,000 (30,896,000)	47,234,000 (35,774,350)	47,113,940 (34,818,370)	36,299,480 (31,767,270)
Print & Ad (P&A) (in $)	15,715,000 (14,620,000)	27,978,000 (13,039,970)	29,342,490 (13,128,040)	22,828,910 (14,316,330)
Weeks in channel (in weeks)	5.31 (2.85)	62.85 (57.22)	90.93 (94.33)	6.85 (3.97)
Inter release time	n/a	87.43 (106.23)	95.20 (108.26)	24.42 (6.06)
Price	n/a	20.08 (4.15)	10.76 (3.43)	n/a

---

[9] Mean followed by standard deviation in parenthesis

Table 1 presents key summary statistics. All variables reported are means across title, in the channel, in a week. For instance, the first row, channel gross, is the total sales of a title in the channel in the week. The average title running in theatres grosses approximately 3.5 million dollars each week of its run, while the average DVD in our dataset, has total sales of approximately 100,000 dollars each week.

The second and third rows are only for secondary channels. Box Office is the cumulative box office (primary channel) gross of a title, now released in the secondary channel. Screen-Weeks is the sum of screens on which the movie showed, over the first 10 weeks of its theatrical run. Budget is the cost of making the film. Print & advertising spending measures expenditure on both advertising as well as the cost of creating prints for distribution. User Rating and Critic's Ratings were described earlier. Weeks in Channel measures how long the movie has been in that channel post release, and Inter Release Time measures the time between primary and secondary channel release.

5.2: Operationalizing the Deterministic Component of Attraction Function

We discuss the mathematical form and variables used in modeling the deterministic component (δ_{mkwy}), first in the primary channel, and then in the sequential channels. In the theatrical channel, Sawhney and Eliashberg (1996) proposed a three parameter gamma model (BOXMOD) for predicting box office revenue. After a meta-analysis of earlier movies, they predict first week, peak and decay of theatrical revenue over weeks. ADZ make the distinction between blockbuster decline (early peak), or a sleeper decline (later peak). In the box office, we extend these two extant models to incorporate more explanatory variables, particularly P & A, user ratings and critics' ratings, improving fit significantly. Suppressing subscripts denoting movie m, in channel k in week w, year y, we write the deterministic component in box office as[10]

$$\delta = p^{\alpha_k}\left[BB:\log(PA)\right]^{\beta_{1k}}\ x^{\beta_{2k}}\, {w_R}^{\left(\beta_{3k}+\beta_{4k}\lg(PA)\right)} e^{\left(\beta_{5k}+\beta_{6k}\lg(PA)\right)w_R} \qquad (7)$$

where p is the price, BB is a set of dummy variables coded using 4 levels of print and advertising expenditure (henceforth P&A) to capture blockbuster status, PA stands for P&A, x is a characteristics vector and w_R is weeks spent in the channel post release.

We assume that the price of a box office ticket remains constant over the length of the dataset. We interact of blockbuster status with P&A to allow for a non linear relationship between gross and P&A. While we observe the P&A for the theatrical channel, we do not observe the marketing mix used by the firm.

For all channels, the characteristics vector x we include the following measures of differentiation among movies: studio dummies, dummies for genre, dummies for animation movies, MPAA ratings, and dummies for the presence of star actors and director. We also use user ratings to measure consumer reported quality and a summary measure of critics' ratings.

In the sequential channels, we interact movie characteristics time spent in channel and inter release time. Suppressing subscripts denoting movie m, in channel k in week w, year y, we write the deterministic component in sequential channels as

---

[10] Comparing fits, we find little difference in using the logarithm of the price and characteristics, and the untransformed variables.

$$
\begin{gathered}
\delta = \left\{ \left[BB : p \right]^{\alpha_{1k}} \; p^{\alpha_{2k} \log(w_R + w_C)} \right\} \left\{ \left[BB : \log(BO) \right]^{\beta_{1k}} \; x_m{}^{\beta_{2k}} \right\} \\
\left\{ w_R{}^{\left(\beta_{3k} + \beta_{4k} \lg(BO)\right)} w_C{}^{\beta_{5k}} \right\} \left\{ e^{\left(\beta_{6k} + \beta_{7k} \lg(BO)\right) w_R + \beta_{8k} w_C + \beta_{9k} w_R w_C} \right\}
\end{gathered}
\tag{8}
$$

where BB is a set of dummy variables coded using 4 levels of box office gross to capture blockbuster status, BO stands for box office gross, PA stands for P&A in the primary channel, w_C is weeks spent (inter release time) between the theatrical and sequential channels and other variables are as described for (8) above.

As explained earlier, we assume that the price of a video rental remains constant over the length of the dataset, and use the mean price in sales. As the price elasticity of a movie in DVD or VHS sales may depend on its box office gross and the time since the movie was released in theaters, we interact blockbuster status with price and time since theatrical release when identifying the price response coefficient.

We observe the P&A for the theatrical channel, but not the marketing mix used by the firm. Lacking similar data at the sequential channels, we use P&A estimates from the box office stage to proxy for spend in promotions at the sequential channel.

For sequential channels, there are other variables included in the x vector. Box office revenue captures the impact of unobservables and is a proxy for market size in secondary channels (Krider and Weinberg, 1998, and Lehman and Weinberg,2000).

We use the number of screens a movie was show in during the first week, and sum over the number of screens a movie was shown in during the first 10 weeks of its run (henceforth Screen-Weeks) to capture distribution effects.[11] Last, we use the ratio of budget to box office gross (henceforth Profitability Index), to differentiate between smaller budget and larger budget films with the same box office gross. Extant forecasting models have not considered multiple channels, and hence do not suggest a relationship between the inter release time (number of weeks between release in the theatrical and sequential channels), w_C, and the decay in revenue over the weeks post release.

The complexity of the underlying cannibilization and dynamic decision making process, and the network effects in evaluating entertainment products would suggest non-linearaties in this relationship. For instance, a longer inter release time may lead to a saturation of the word of mouth, attracting more consumers in earlier weeks, and then showing faster decay post release? Alternatively movie with shorter inter release times may attract more customers in earlier weeks because of the heavy advertising in the box office channel, and then show faster decay post release?

5.3: Results and Model Comparisons

In this subsection, first we discuss the coefficients estimated. Second, we present our

---

[11] Major studio typically release movies first in the theatres, and then for sale in VHS/DVD format and in rental stores. While the primary to secondary distribution channel gap has shrunk in recent times (Luan and Sudhir, 2007), the mean duration in our dataset is 24 weeks. Thus, it is safe to assume that few movies remained in theatres for 6 months. Hence, cannibalization of sales across sequential distribution channels will primarily occur through forward looking conjectures of theatrical consumers, and not due to direct substitution of a movie at the theatre with the same movie at a rental store.

predictions and compare them with other models. Table 2 reports coefficients estimated for the attraction function in each channel. Table 3 reports elasticities estimated at mean levels of variables.

Table 2: Coefficients of Market Attraction[12]

	Box Office	***DVD***	***VHS***	***Rentals***
log(Budget)	0.208 ** (0.031)	-0.037 ** (0.011)	-0.006 (0.012)	0.097 ** (0.014)
log(Screens in Week 1)	n/a	0.115 ** (0.007)	0.073 ** (0.007)	-0.026 ** (0.007)
log(Screen Weeks)	n/a	0.011 (0.017)	0.058 ** (0.019)	0.116 ** (0.013)
log(Print and Advertising) (P&A)	n/a	-0.155 ** (0.014)	-0.227 ** (0.015)	0.016 (0.015)
Profitability Index	n/a	-0.00002* (0.00000)	- 0.00005** (0.00002)	0.000002 (0.000004)

---

[12] We do not report coefficients on animation, genre and rating. B1, BB2, BB3 and BB4 are the blockbuster status, in ascending order of P&A for primary channels and Box Office Gross for secondary channels. **Significance codes: 0.001 '**' 0.01 '*' 0.05 '$^+$'**

Table 2 (continued)

log(User Rating)	1.110 ** (0.149)	2.06 ** (0.053)	0.633 ** (0.048)	0.565 ** (0.063)
log(Critics' Ratings)	0.377 ** (0.044)	-0.180 ** (0.015)	-0.004 (0.014)	-0.116 ** (0.017)
Weeks since release (WR)	0.388 ** (0.122)	0.019 ** (0.001)	0.006 ** (0.001)	-0.215 ** (0.015)
Inter Release Time (WB)	n/a	0.002 ** (0.000)	0.005 ** (0.000)	-0.034 ** (0.006)
WR*WB	n/a	- 0.000003* (0.000001)	-0.00002** (0.00000)	0.001 ** (0.000)
log(WR)	1.610 ** (0.502)	-0.725 ** (0.034)	0.282 ** (0.033)	0.352 ** (0.063)
log(WB)	n/a	-0.024 (0.023)	-0.195 ** (0.022)	0.462 ** (0.165)
Star Actor	-0.115 * (0.052)	-0.009 (0.017)	-0.002 (0.017)	0.025 (0.018)
Star Director	0.107 (0.100)	0.020 (0.019)	-0.162 ** (0.020)	-0.061 * (0.029)
BB1* Price	n/a	-0.047 ** (0.004)	0.051 ** (0.005)	n/a
BB2* Price	n/a	-0.017 ** (0.005)	0.047 ** (0.005)	n/a

Table 2 (continued)

BB3* Price	n/a	0.091 ** (0.015)	0.035 ** (0.012)	n/a
BB4*Price	n/a	-0.124 $^{+}$ (0.064)	0.003 (0.055)	n/a
(WR+ WB)* log(Price)	n/a	-0.116 ** (0.012)	-0.215 ** (0.010)	n/a
BB1*log(P&A) for primary BB1*log(Box office) for secondary channels	1.250 ** (0.044)	0.786 ** (0.026)	0.961 ** (0.028)	0.582 ** (0.019)
BB2*log(P&A) for primary BB2*log(Box office) for secondary channels	1.300 ** (0.041)	0.694 ** (0.031)	1.020 ** (0.030)	0.526 ** (0.018)
BB3*log(P&A) for primary BB3*log(Box office) for secondary channels	1.340 ** (0.041)	0.126 $^{+}$ (0.067)	1.020 ** (0.038)	0.550 ** (0.022)
BB4*log(P&A) for primary BB4*log(Box office) for secondary channels	1.350 ** (0.042)	1.090 ** (0.236)	1.100 ** (0.177)	0.464 ** (0.028)
WR*log(P&A) for primary WR*log(Box Office) for secondary	-0.073 ** (0.014)	-0.002 ** (0.000)	0.001 ** (0.000)	-0.031 ** (0.004)

Table 3: Elasticity of Market Attraction[13]

	Box Office	*DVD Sales*	*VHS Sales*	*Rentals*
Price	n/a	0.607	-1.09	n/a
Box Office Gross	n/a	0.436	-0.006	0.530
Screen Weeks	n/a	0.011	0.058	0.116
Screens 1st Week	n/a	0.120	0.073	-0.026
Budget	0.208	-0.040	-0.006	0.097
Print & Advertising	0.576	-0.160	-0.227	0.016
Profitability Index	n/a	-0.00002	-0.00005	0.000002
Time Since Release	-6.21	-3.03	-1.63	-4.92
Inter Release Time	n/a	1.62	0.117	-0.164
User Ratings	1.11	2.06	0.633	0.565
Critics' Ratings	0.377	-0.180	-0.004	-0.116

We find that market attraction is well predicted by the print and ad spending of a movie, the user ratings and critics' ratings of a movie, all of which are positively correlated with larger box office revenues. Larger number of weeks since theatrical release significantly decreases the attractiveness of the movie.[14] As in Lehmann and Weinberg (2000), we find that performance measure from the primary channels help

---

[13] Elasticity computed using dataset means, and accounts for all significant parameters associated with the variable.

[14] We try different time specifications and do not see a difference in fit across different specifications, including higher order polynomials of time spent in channel.

improve fit and prediction in the secondary channels. Larger box office gross and greater screen-weeks exposure leads to larger secondary channel attractiveness. Similar to Luan and Sudhir (2007), we find that longer inter release times between channels, decreases market attractiveness.

We compare three models that have the same deterministic specification for market attraction. Similar to ADZ, first write the generalized gamma formulation of BOXMOD as:

$$S_{mkwy} = \eta_i t^{\gamma_i / \beta_i} e^{1-t / \beta_i}$$

One can rewrite BOXMOD as:

$$S_{mkwy} = \beta_{1i} t^{\beta_{2i}} e^{\beta_{3i} t} \qquad (9)$$

As the original model does not restrict the three parameters of the gamma function $(\eta_i, \gamma_i, \beta_i)$, (9) is a re-parameterization of the original formulation. Comparing (9) with(5), our model has to additionally account for the inter release time and its interaction with the time spent in the channel. Hence, when modeling secondary channels, we modify BOXMOD to include time spent in channel and inter release time (t_C, t_{IR} respectively):

$$S_{mkwy} = \beta_{1i} t_C^{\beta_{2i}} t_{IR}^{\beta_{3i}} e^{\beta_{4i} t_C + \beta_{5i} t_{IR} + \beta_{6i} t_C t_{IR}} \qquad (10)$$

To reduce computational complexity, BOXMOD can be estimated using a hierarchical two step approach: the three parameters for the gamma function are fitted separately for each movie and then the estimated parameters projected onto the observables of the movie. Substituting the linear model for each parameter and then taking logs on both side, yields the efficient estimator. Noting that (7) and (8) are the best fits[15] for this class of models and substituting in (9) and (10), we get:

$$\ln(S_{mkwy}^{M1}) = 1 + \ln(\delta_{mkwy}) + \xi_{mkwy} \tag{11}$$

In the second model, we include a weekly seasonality dummy to allow for seasonal trends, and a yearly dummy to control for technology trends:

$$\ln(S_{mkwy}^{M2}) = t_{ky} + t_{kw} + \ln(\delta_{mkwy}) + \xi_{mkwy} \tag{12}$$

(11) and (12) are estimated by OLS. The third model is the efficient estimator model. This utilizes the methodology described above to predict both market shares and the aggregate rental revenue. While the first and second model outperform our model in the box office validation sample, our model has the lowest rMSE in all secondary channel validation samples. In DVD sales, VHS sales and Rentals, our estimator shows 2%, 4% and 6% lower rMSE respectively.

There are two important reasons why this might be so. First, primary channels are less likely to be prone to cannibalization or complementarity from other channels when

---

[15] White's correction, clustering errors by time spent in channel, inter release time, and/or week of observation, do not improve fit.

compared to secondary channels. Our model places heavier demand on data to estimate this flexible competition model (within and) across channels, and therefore appears to not do as well in primary channels where this is less of a concern. Second, unobservables play a larger role in the box office channel model than they do in the secondary channel model. Recall that in the secondary channels, the presence of box office gross controls for many of these. Hence, all three forecasting models perform very poorly in the primary channel.

Table 4: rMSE of models

	BOXMOD equivalent	***Seasonal BOXMOD equivalent***	***Our model***
Box Office (training)	1.46E+06	1.56E+06	5.90E+05
Box Office (validation)	3.99E+05	4.59E+05	8.43E+05
DVD Sales (training)	1.21E+04	1.24E+04	1.27E+04
DVD Sales (validation)	4.27E+04	3.27E+04	3.21E+04
VHS Sales (training)	7.76E+03	7.91E+03	7.64E+03
VHS Sales (validation)	5.74E+04	5.62E+04	5.38E+04
Rentals (training)	6.28E+05	4.66E+05	4.38E+05
Rentals (validation)	5.10E+05	4.58E+05	4.29E+05

6. Conclusion

We describe, estimate and then benchmark a model to predict weekly rental revenue by title. Our model incorporates both seasonal demand variation and the market effects of better movies being released in periods of peak sales. To ensure managerial relevance, the model only utilizes data from available months prior to the week of interest. We find that prior channel performance helps predict future performance. However, we cannot and do not differentiate between causality and correlation. A larger box office gross might potentially lead to more word of mouth and hence more future revenues, or might simply indicate movies of higher quality. Greater distribution (screen-weeks) might indicate a longer runs (better quality) or might capture increased cannibalization across channels. Allowing for a general model of competition within and across channel keeps the model flexible without separating these forces.

Additionally, we find that movies show decay in market share with inter-release and time spent in the channel, and that this effect is highly nonlinear. We find the interaction of these variables remains critical to good forecasts. The non-linearity of time trends in the market attractiveness function process hints both at the complexity of the underlying cannibilization and dynamic decision making process, and the network effects in evaluating entertainment products. This is an interesting avenue for future research.

Our model has some drawbacks. First, we assume inviolate channel boundaries. As different channels begin to overlap both in product and the timing of release, the model will be affected by endogeneity concerns and network effects. Second, our model does not include other sources of entertainment that compete with movies. Not including competitive sources of entertainment, such as television shows when modeling rentals, increases forecast error. While we control for seasonal trends, including major entertainment events such as the Super Bowl, might improve predictive capabilities. Finally, in markets where long stable sales histories are available to marketing managers, our method will have limited usability. Using brand specific parameters and past sales should provide superior forecasts, but are unavailable or inappropriate in rapidly changing industries.

BIBLIOGRAPHY

Ainslie, Andrew, Xavier Dreze, and Fred Zufryden (2005), "Modeling Movie Life Cycles and Market Share," *Marketing Science*, 24 (3), 508-17.

Chiou, Lesley (2007), "The Timing of Movie Releases: Evidence from the Home Video Industry," *working paper*, Occidental College, CA, USA.

Cooper, Lee G. and Masao Nakanishi (1988), *Market-Share Analysis: Evaluating Competitive Marketing Effectiveness*. Boston: Kluwer Academic Publishers.

Einav, Liran (2003), "Not All Rivals Look Alike: Estimating an Equilibrium Model of the Release Date Timing Game," *working paper*, Stanford University, CA, USA.

---- (2007), "Seasonality in the U.S. Motion Picture Industry," *RAND Journal of Economics*, 38(1) (forthcoming).

Elberse, Anita (2006), "The Power of Stars: Do Star Actors Drive the Success of Movies?," *working paper*, Harvard University, MA, USA.

Elberse, Anita and Jehoshua Eliashberg (2003), "Demand and Supply Dynamics for Sequentially Released Products in International Markets: The Case of Motion Pictures," *Marketing Science*, 22 (3), 329-54.

Elberse, Anita and Felix Oberholzer-Gee (2007), "Superstars and Underdogs: An Examination of the Long Tail Phenomenon in Video Sales," *working paper*, Harvard University, MA, USA.

Eliashberg, Jehoshua, Anita Elberse, and Mark A. A. M. Leenders (2006), "The Motion Picture Industry: Critical Issues in Practice, Current Research, and New Research Directions," *Marketing Science*, 25 (6), 638-61.

Eliashberg, Jehoshua, Jedid-Jah Jonker, Mohanbir S. Sawhney, and Berend Wierenga (2000), "MOVIEMOD: An Implementable Decision-Support System for Prerelease Market Evaluation of Motion Pictures," *Marketing Science*, 19 (3), 226-43.

Epstein, E. J. (2005), "*The Big Picture: The New Logic of Money and Power in Hollywood*" Random House.

Hennig-Thurau, Thorsten, Victor Henning, Henrik Sattler, Felix Eggers, and Mark B. Houston (2007), "The Last Picture Show? Timing and Order of Movie Distribution Channels," Journal of Marketing, 71 (4), 63-83.

Hettrick, S (2000), "Raise the Rent," Vol. 2007: *Video Business*.

Krider, Robert E. and Charles B. Weinberg (1998), "Competitive Dynamics and the Introduction of New Products: The Motion Picture Timing Game," *Journal of Marketing*, 35 (1), 1-15.

Lee, Jonathan, Peter Boatwright, and Wagner A. Kamakura (2003), "A Bayesian Model for Prelaunch Sales Forecasting of Recorded Music," *Management Science*, 49 (2), 179-96.

Lehmann, Donald R. and Charles B. Weinberg (2000), "Sales through Sequential Distribution Channels: An Application to Movies and Videos," *Journal of Marketing*, 64 (3), 18-33.

Luan, Jackie and K Sudhir (2007), "Optimal Inter-Release Time Between Sequential Products: Application to Theatrical Movies and DVDs," *working paper*, Dartmouth College, NH, USA.

Mortimer, Julie H. (2004), “Price Discrimination and Copyright Law: Evidence from the Introduction of DVDs”, *working paper*, SSRN.

Nakanishi, M. and L. G. Cooper (1982), "Simplified Estimation Procedures for MCI Models," *Marketing Science*, 1 (3), 314-22.

Neelamegham, R. and P. Chintagunta (1999), "A Bayesian Model to Forecast New Product Performance in Domestic and International Markets," *Marketing Science*, 18 (2), 115-36.

Prasad, A, B Bronnenberg, and V Mahajan (2004), "Product Entry Timing in Dual Distribution Channels: The Case of the Movie Industry," *Review of Marketing Science*, 2 (2004), Article 4.

Radas, Sonja and Steve M. Shugan (1998), "Seasonal Marketing and Timing New Product Introductions," *Journal of Marketing Research*, 35 (3), 296-315.

Sawhney, Mohanbir S. and Jehoshua Eliashberg (1996), "A Parsimonious Model for Forecasting Gross Box-Office Revenues of Motion Pictures," *Marketing Science*, 15 (2), 113-31.

Shugan, Steve (1998), "Forecasting Failure and Success of New Films," *working paper*, University of Florida, FL, USA.

Smith, Michael D. and Rahul Telang (2006), “Piracy or Promotion? The Impact of Broadband Internet Penetration on DVD Sales,” *working paper*, SSRN.

Swami, Sanjeev, Jehoshua Eliashberg, and Charles B. Weinberg (1999), "SilverScreener: A Modeling Approach to Movie Screens Management," *Marketing Science*, 18 (3), 352-72.

Ulmer, James (2000), *Hollywood Hot List*. New York: St Martin's Griffin.

Vogel, Harold L. (2004), *Entertainment Industry Economics: A Guide for Financial Analysis* (6th ed.). Cambridge ; New York: Cambridge University Press.

CHAPTER 2

COMPETITIVE DYNAMICS OF DVD RELEASE TIMING & PRICING[16]

1. Introduction

In the U.S. motion picture industry, DVDs have increasingly become a major source of revenue for movie studios with total DVD sales growing from $1.9 billion in 2000 to $16.3 billion in 2005. The release date and price of a DVD (or title, as it is called in the industry) affect profits because weekly DVD sales vary dramatically with peak weekly sales being seven-fold non-peak sales, and release prices for DVDs varying between $5 and $35 across titles, in the time period of our study, 2000-2005. Therefore, release timing and prices of DVD are important strategic variables. The focus of this study is the competitive dynamics of the setting of these two variables in the U.S. DVD industry in this period. Substantively our objective is to study the optimality of observed decisions. We investigate changes in DVD revenue and release costs due to time varying title and industry forces, and build managerial decision tools to predict future demand and competition levels.

When deciding DVD release strategies, studios must consider the following. First,

---

[16] This essay is co-authored with Vrinda Kadiyali, a Professor of Marketing and Economics at the S. C. Johnson Graduate School of Management, Cornell University, 385 Sage Hall, Ithaca, NY 14853. We thank seminar participants at Cornell University for comments. We also thank Nielsen EDI, Paul Kagan and Associates, Nielsen VideoScan, and Home Media Retailing for providing data for this study.

there is substantial difference in industry demand between peak demand (weeks with highest aggregate industry sales) and non-peak demand weeks. Cēterīs pāribus, any DVD will obtain higher revenues in a peak-demand week than in a non-peak week. Second, despite peak periods having the potential for higher revenues, any one DVD might not realize higher revenue. This is because the higher revenue potential in a peak week is likely to attract more titles that also anticipate higher revenues. Higher competition in peak weeks is likely to reduce sales of the title, and might result in lower prices, resulting in lower profits. As more titles enter peak demand weeks, release costs in these weeks might be higher. For example, promotional allocations and slotting fees to retailers and the costs of advertising media might rise in weeks of peak demand. In addition, titles have their "shelf-life" clocks running from the time that they exit the theatrical channel. The third factor to consider is that deferring a DVD release (e.g. to a non-peak demand week) long after the movie's theatrical run reduces the potential sales of the title in the channel. Studios must therefore make appropriate strategic choices of DVD release timing and pricing. Given these forces affecting release date and price decisions of any one title, how does the competitive equilibrium of release dates and prices evolve? There are three factors that complicate our modeling the equilibrium.

First, key information describing the evolution of the industry is missing in our data. We know that studios consider various release date options as evidenced by their intermediate release announcements.[17] That is, studios can and do announce they will release on a certain date and price but the final date and price often turns out to be different. Various competitive dynamics are possible in these release date announcements. For example, a title that was successful in the box-office might announce to pre-empt another successful movie of the same genre from releasing the same weekend. At the same time, seeing the release announcement by a high-revenue potential title might persuade a smaller title of the same genre to release in the same week, in the hope of free-riding off promotional monies spent by the larger title, especially if the peak demand favors the particular genre of movies. This might in turn deter another smaller movie of the same genre from announcing the same release date. Therefore, these announcements can serve to preempt, or coordinate release dates with rivals.

We do not observe the interim release announcements in our dataset, and only observe the final release date and time. We build a model that allows for the effect of the missing data on agent actions when finding the profit function, and show estimates from the model converge to the traditional full information MPNE estimator. An alternative choice is to build a model of the release timing and pricing decision taken at the last period prior to release, ignoring interim announcements. Such a dynamic game is fully observed and can be estimated using a Markov Perfect Nash Equilibrium

---

[17] Cached web-pages accessed using the internet archive confirm the existence of a steadily changing set of announcements. However, we have been unable to find a data source for past DVD announcements of studios.

(MPNE) frame work. However, in general, a model ignoring the strategic variables of the game (release timing and pricing announcements) may be misspecified, and lead to biased profit function estimates.

Hence, our paper contributes methodologically by describing an estimator for agent payoffs in a dynamic model with censored (or missing) information. Marketing researchers are often faced with datasets on forward looking firms in which a key strategic variable is unobserved or censored for a part or the entirety of the dataset. For example, firms might scout several locations before choosing a final location, or make a capacity decision and then alter it before reaching the final choice. A typical MPNE estimation needs complete data on current and future state vectors and actions taken by agents to identify the transition matrix and enable use of nested fixed point algorithms. Our model is general enough to be applied to other industries with censored data.

Second, the seasonality of demand and competition leads to time-varying strategy selection rules for firms. Each title is likely to have a release timing and pricing strategy that varies by week. For instance, a title may be more likely to release a DVD title if future periods have decreased demand, than if future periods have increased demand; the seasonality of demand leads to seasonality in the set of entrants. Also, in this market, there was an eight-fold increase in demand between 2000 and 2005. The seasonality of demand itself changed over time (DVDs were gifted not just for Christmas but also increasingly for Mother’s Day, or graduation, or children’s birthdays). Hence, studio profits varied over weeks and over years. In our model, this implies that the states of the world do not follow a first-order stationary (homogenous) Markov process, as is typically assumed in MPNE models.

Third, it is plausible that our setting has multiple equilibria. Release timing (and pricing) strategies in our model may be both the traditional strategic substitutes but also strategic complements. For example, a movie with small theatrical revenues might view a larger movie as a strategic complement if the larger movie can drive traffic to DVD stores. Two larger movies may see their timing strategies as strategic substitutes if the market stealing effect dominates. Allowing for strategic complementarity and substitutability brings about the possibility of multiple equilibria, a situation not tractable in extant dynamic model estimation methodologies.

We describe a novel estimator for dynamic games based on a partial information model (i.e. where the announcement vector is unobservable). Our model is analogous to the Oblivious Equilibrium (OE) model (Weintraub, Benkard and Roy, 2007). OE assumes that an “agent” (studio in our model) in a period is “oblivious” to the current states of others, and instead holds beliefs (distribution) over candidate states possible in the period. Similar to OE, in our partial information model, studios are oblivious to release announcements of other titles in past periods, and instead have beliefs that reflect the Perfect Bayesian distribution over possible release announcements. In appendix 1.2 we show that payoffs estimated in the partial information approach converge to those in a full information model. While our partial information model is unable to identify the drivers of an interim release announcement in the game, the model is able to identify the relevant components of studio profit functions and hence the trade-offs made by the studios between forces.

The estimation methodology proposed in OE is not appropriate for our setting. Extant MPNE models, including OE, identify model primitives by using the revealed preferences of a firm in a period. The methods identify drivers of profit functions by comparing observed choices with computed best responses. The missing announcements prevent us from using extant methods. We show that we can recast the equilibrium condition, as being a fixed point in best responses of agents across multiple periods; theorem (T1) shows that the optimal release announcement strategy of a firm leads to maximum future profits from release, and hence an optimal final release timing and pricing schedule.

We show that a unique industry evolution pathway is consistent with our partial information model. We estimate a market outcome function that describes sales in a period, as a function of the time since theatrical and DVD release, seasonal demand and competition. Employing a logit formulation of market shares, we generate a sufficient statistic to approximate the evolution of the industry equilibrium which accounts for differences amongst titles and studios. We forecast industry evolution using the statistic, and compute payoffs to studios from different release timing and pricing choices. Last we maximize the quasi-likelihood of observed strategy choices, consistent with the other steps of the estimation, to measure the seasonal differences in release costs which rationalize observed actions.

Our estimates of DVD market share are similar to prior findings.[18] We find that net release costs are smaller for movies that were more successful in the box office, indicating that blockbusters both make more money on DVD, and face lower release costs than other titles. Increasing the inter-release time (time between theatrical and

---

[18] Due to the high computational load of the estimation algorithm, results discussed are preliminary and have been estimated on data from 2000 to 2002. We will shortly estimate the model on the entire 2000-2005 data.

DVD release) both reduces the sales potential of the title, and increases release costs. The model predicts competition levels in the industry with reasonable accuracy up to 30 weeks into the future, providing a decision support tool for marketing managers in studios. Within sample, our model shows a good fit with higher prediction accuracy for future release dates and prices of DVDs than alternative model specifications. Finally, our policy simulations investigate how optimal Theater-to-DVD windows depend on seasonal demand, competition and release cost variation. The simulations provide a decision support tool for marketing managers in the entertainment industry, predicting changes in DVD release strategies should there be changes in the industry landscape (market expansion, change in seasonality, change in release costs, etc.).

Our model generalizes the MPNE frame work used to study strategic decisions by multiple forward looking firms, and is applicable to competitive industries with time-varying payoffs or with frequent entry of new products. Examples include technology products, other entertainment products like music, and fashion products. And as described earlier, the framework can be used in industries where researchers are unable to obtain data on firm actions/state space, and need to model using a censored dataset.

2. Conceptual Overview of Equilibrium Forces

As mentioned in the introduction, three forces affect the equilibrium. First, weeks of peak demand increase sales. Second, intense competition in weeks of peak demand reduces market share and lowers margins. Relatedly, peak weeks might have higher costs of release. Third, deferring a DVD release (e.g. to a non-peak demand week in order to avoid competition) to a week long past the movie’s theatrical run reduces the potential sales of the title in the channel. In this section we present a conceptual overview of our model, reviewing relevant literature and discussing each of the forces.

2.1: Temporal Variation in Demand

As mentioned in the introduction, demand for DVDs is time varying, affecting the optimal release timing and pricing of titles. We first discuss how existing literature has approached seasonality, and then its impact on our model.

In the economics literature, Einav (2007) and Chiou (2007) study the impact of seasonality on the demand for movies in primary and sequential channels respectively. Einav (2007) presents an empirical analysis of theatrical revenue in the U.S. movie industry, studying both seasonality and competition. Because his model is for theatrical releases, his research question does not include declining potential sales with deferral of entry dates. Chiou (2007) includes the effect of deferral of release dates, controlling for the endogeneity of release date selection, but does not model the process for the evolution of dates and prices. Both papers find strong evidence of seasonal changes in demand, and that firms' account for seasonality in their strategic choices.

The marketing literature has found that revenue in the movie theater, over the same weeks in different years, can be predicted. Krider and Weinberg (1998) discuss competition when faced with seasonal demand variations. Radas and Shugan (1998) outline an approach for including seasonal trends in estimating demand curves by taking a transformation of time. Luan and Sudhir (2007) model the effect on box office revenues of the theater-to-DVD window.

Seasonal demand (and seasonal costs, as we will discuss below) cause seasonal variation in payoffs. Modeling these seasonal payoffs in the framework of dynamic games poses a problem. While dynamic games have been studied for a few years now (e.g. see Rust 1987, Ericson and Pakes 1995), most researchers have typically focused on mature and stable industries. Firms in these industry use time invariant strategy selection criteria, allowing the researcher to assume that the states of the world follow a stationary, first order Markov process. The stationary Markov process leads to the solution concept of a stationary MPNE, using the implicit assumption that profits are only a function of strategic decisions of firms, with no exogenous change in industry profits due to macroeconomic forces or technological change.

The DVD industry in our period of interest showed a rapid increase in sales from $2 billion to $16 billion in 6 years. The growth in sales did not occur symmetrically over various weeks in a calendar year, nor across various calendar years. For instance, the largest growth in DVD sales occurred in the Christmas holidays. Thus, neither over weeks of the same year nor in the same season across multiple years, were revenues and costs, hence profits and firm strategy selection criteria, constant as assumed in the stationary MPNE model.

We utilize the Ericson and Pakes (1995) framework, and add to the MPNE literature by describing a non-stationary (time varying) MPNE model. Our model relaxes the assumption of time homogeneity for MPNE. In appendix 1.2, we discuss the relationship between the non-stationary and stationary MPNE in greater detail. In general, the theory does not guide us on how non-linearities of response functions may translate to decision making rules. We show that the non-stationary MPNE requires a “sufficient statistic” vector describing changes over time, which is essential for identification (see assumption A7 in appendix 1.2 for further details).

Assuming the sufficient statistic for change over time, a non-stationary MPNE can be estimated using extant methods for stationary MPNE. However, the non-homogeneity of model primitives substantially increases the data requirements for estimation and decreases the rate of asymptotic convergence in extant models. In particular, the effect of seasonal change enters non-linearly in the continuation values of the MPNE (see Pakes and McGuire, 2001 for a discussion on calculating continuation values). In our partial (limited) information model, we impute changes in payoffs, through a demand function. The non-linear effects of seasonality are partially accounted for in the demand function. Hence the effect of seasonality enters linearly when calculating continuation values, decreasing computational and data requirements for estimation. We discuss the estimation algorithm used in greater detail in section 4.4.

2.2: Temporal Variation in Competition

Seasonal variations in demand are likely to result in seasonally varying levels of competition (Figure 2). Both revenues and costs seasonally due to changes in competition, leading to seasonally varying profitability. We now turn to more detailed discussion of this phenomenon.

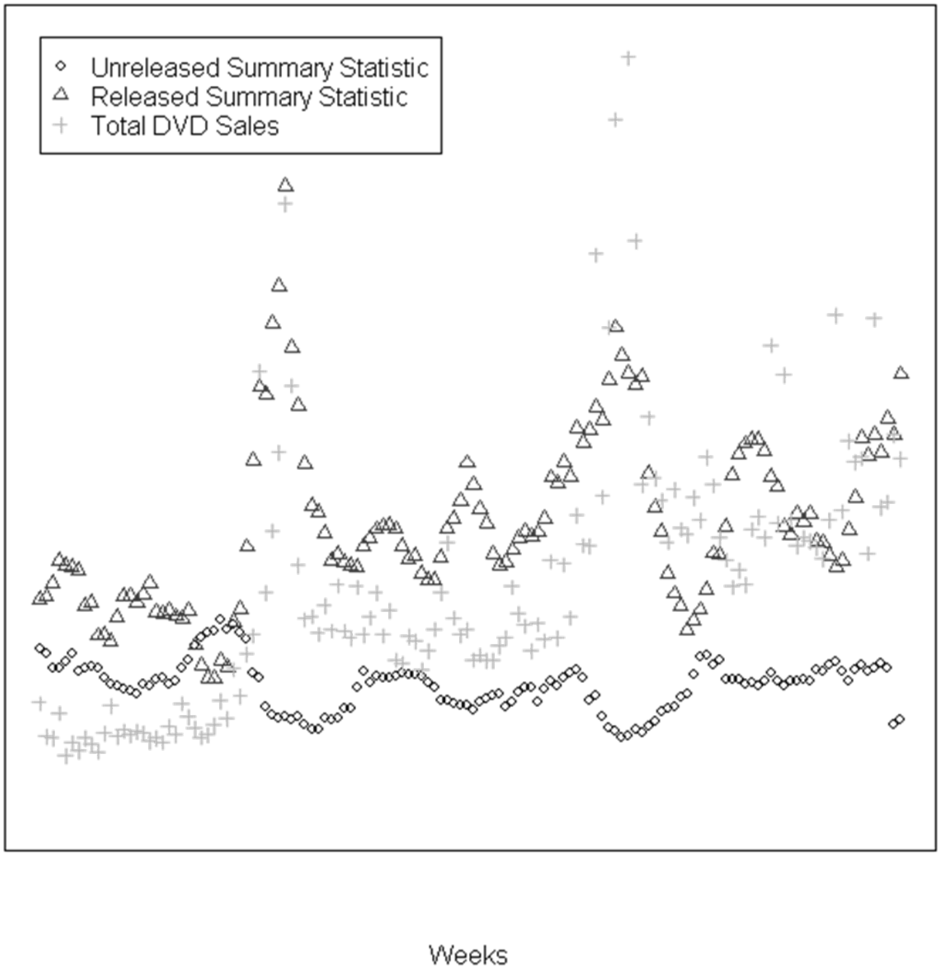

Figure 2: Total DVD Sales and New Releases

Consider first papers on the impact of competition in movies. Swami, Eliashberg, and Weinberg (1999) study multiplex screen allocation decisions and formulate a model to optimize exhibitor scheduling. Ainslie, Dreze and Zufryden (2007) build on the BOXMOD model and study the lifecycle of a movie at the box office, measuring competition within a channel. Einav (2003) models the release timing game in theatrical channels as a sequential game of imperfect information. Foutz-Zhang and Kadiyali (2007) model the release timing game in the theatrical channel and find that pre-announcements of release dates for movies serve a strategic function to deter entry into holiday weeks.

Studios can mitigate competition by making (interim and final) release announcements[19], which allow the studios to compete and cooperate on release schedules. There may be early/late mover advantages to announcing. For instance, as mentioned in the introduction, a title with large theatrical revenue might announce that it is releasing in a high-demand week. This might deter a title of the same genre (and that had a worse theatrical performance) from announcing a release in the same week. Or the similar title might announce the same release date, but set a lower price to undercut the first title. Smaller theatrical-revenue titles might prefer to announce a non-peak demand week where competition is less intense.

Allowing for strategic complementarity and substitutability brings about the possibility of multiple equilibriums Announcements in our model carry the ability to both coordinate and/or pre-empt release timings and pricing. Depending on which pair of competitive interactions are being studied and depending on the revenue potential of the week, strategies in our model may be both strategic substitutes and/or strategic complements. For instance, implicit price collusion due to grim Nash reversion, can be sustained in the model through the presence of a punishment state (low release price), with the Markov kernel capturing the probability of agents entering and exiting the

---

[19] The release announcements are not made in consumer outlets and hence do not serve to inform or influence potential consumers. They are reported in industry websites like videoeta.com. Although complete cached data on the history of these announcements is not available, there is strong evidence of titles changing announced release dates and prices.

punishment regime. Strategies in this model are strategic complements in the collusive regime, and strategic substitutes in the punishment regime.

While recent advances in game theoretical modeling have lead to estimation techniques for static models with multiple equilibriums, to our knowledge this is the first paper to allow multiple equilibriums to be played in the data. The non-stationary MPNE model allows "equilibrium switches" in a single path of play, as the Markov kernel is not restricted across multiple periods. That is, in each period, agents can choose to play strategies leading to a different equilibrium. Optimal strategies played in the data, are not restricted to being the same across different seasons and different years. For instance, equilibriums played over the summer may be different from equilibriums resulting in the holiday season. Our method describes a consistent estimator for payoffs in a dynamic multi-agent game, accounting for randomization between equilibriums. A caveat is that we do not distinguish between potential equilibriums and cannot identify the probability of choosing a given equilibrium in a period. The Perfect Bayesian evolution of the industry is the result of the equilibrium conditional on equilibrium choice, and the equilibrium choice probabilities.

An important point to consider is whether these announcements reflect actual intentions of titles' release timing and prices, or whether they are strategic lies or simply cheap talk. In formulating a competitive dynamics model for release timing and price, we do not require firms to take decisions influenced by competitor announcements. For instance, if announcements are cheap talk, then in our model the Markov density will reflect the lack of information in the announcements. On the other hand, it is possible that announcements are not cheap talk, and are instead costly commitments to particular strategies. In this case, our model provides consistent estimates of parameters of DVD profits, while allowing for the competitive and cooperative incentives of the release timing game. Thus, the model nests cases where announcements do not shape release and pricing schedules observed, while allowing for the strategic importance of these decisions.

In the literature, entry and post-entry competition have mostly been modeled separately with extant papers on entry and entry timing, typically using two period static entry models (Mazzeo, 2002) in industries where the researcher observes the release of new products (Einav, 2003), or the entry of a firm in multiple locations (Seim, 2006), but not sales post release. A notable exception is Ellickson and Mishra (2007), who use market outcome models to enrich the description of payoffs in a static game.

Our model draws on both product choice and dynamic entry models to recover studio- and season-specific release costs, providing a richer description of the industry. Prior models of release timing in movies have focused exclusively on the seasonality of consumer demand and competition in the week as a source for time-varying profits, and the explanation for observed seasonality in studio actions (Einav, 2003). There are several reasons for expecting unobserved releases costs to vary over the course of the year. 75% of studios' marketing budgets on average are dedicated to broadcast media (Galloway, 2004), where the cost of advertising varies over the course of the year. DVD release costs may vary as a function of the total sales in the week of release, as channel partners may be able to demand a better share of profits in weeks of high demand. The increased competition between DVDs may spill over into within-store promotions (e.g. end cap displays) and other retailer promotional resources. Lastly, the growth in total DVD sales may lead to a change in release costs and retailer margins. We measure how release costs vary over studios, and over time, and incorporate the effect of changing release costs when conducting counterfactuals and simulations, and compare model fits with and without release costs.

2.3: Perishability

We mentioned in the introduction that the lag between a title's theatrical run and its DVD release (which we term inter-release timing) is likely to have implications for the title's profitability. Movies lose appeal as they spend longer times between sequential distribution channels, a demand feature we call inter-release perishability. Additionally, movies lose appeal after release in a channel, a feature we call within-channel perishability. Below, we discuss how researchers have modeled perishability and explain our conceptualization in our competitive equilibrium framework.

Luan and Sudhir (2007) model the impact of cannibalization of box office revenues by sales and rentals of DVDs, accounting for forward looking behavior of the consumer at the theatre. While cannibalization of theatrical sales provides an incentive for a studio to increase inter-release times, the need to release a movie fresh in the minds of a consumer provides an incentive to decrease inter-release times. Thus they study the underlying tradeoffs between earlier and delayed releases in secondary channels on theatrical revenues.

Three papers study optimal firm actions to maximize revenue in the sequential channel, when considering the perishability of a title. These papers suggest that word of mouth, advertising wear-out effects, and network effects can explain perishability. Hennig-Thurau et al. (2007) use conjoint data to study the effect of different configurations of sequential distributional channels on studio profitability, optimizing release timings across channels. Lehmann and Weinberg (2000) develop a model of the optimal time to enter video rentals for a movie, accounting for the cannibalization of sales from theatrical release. Prasad, Bronnenberg, and Mahajan (2004) use an analytical model to study the effect of consumer expectations on the optimality of the timing decision. In their model, the duration between theatrical and DVD releases of earlier movies, shapes the beliefs of a forward looking-customer for a new movie. The studio's decision depends on current beliefs, making it profitable to deviate from the industry standard, and release early. In each model, firm actions in a title are studied in isolation of the presence of other titles, and of seasonality.

Movies in theaters only exhibit within-channel perishability, Ainslie et al (2007) separate revenue patterns in movie theaters into blockbuster patterns and sleeper patterns. Blockbusters peak early in the first weeks post release, and then decline in revenue. Sleepers peak later than blockbusters, and subsequently decline in revenue. Revenue patterns in DVDs are more complex as DVDs exhibit both forms of perishability. A consumer's dynamic decision making process and the network effects in evaluating entertainment products lead to non-linearaties in the relationship between inter-release perishability and within channel perishability. For instance, longer inter-release times may lead to a saturation of the word of mouth, attracting more consumers in earlier weeks, and then showing faster decay post release. Alternatively shorter inter-release times may attract more customers in earlier weeks due to advertising in the box office channel, and then show faster decay post release. In our paper, we do not separate effects leading to wear-out and instead adopt a flexible 3-parameter gamma specification in the demand formulation.

Managerially, our model can be used to understand the effect of shorter/longer average theater to DVD windows on DVD profits, by simulating the competitive equilibrium in the industry for different DVD release strategies. Extant research on the window between channels has focused on the change in revenue in a particular title, without accounting for seasonality and competition (Luan and Sudhir, 2007). Our paper focuses on the competitive and seasonal aspects of release timing, and their effect on DVD profits. The results from simulations in our model may differ from the inference in models that ignore seasonality and competition. For instance, Indiana Jones 4 was released in theaters on May 22, 2008. A model that ignores seasonality may find that shorter theater to DVD windows are optimal and hence suggest a date in August or September 2008 for Indiana Jones 4. However, seasonality in different channels suggests that optimal decisions differ based on time of theatrical release. Indiana Jones 4 may be better served by waiting for the Christmas Holidays, postponing the DVD release by a period longer than the industry average theater to DVD window.

3. Data and Model

3.1: Descriptive Statistics

Our data comprises release dates, quantities and prices of titles per week after release, as well as title-specific descriptors (e.g. box office revenue, etc.) for all DVDs released in the United States between 2000 and 2005. We describe below the sources of these data and issues with them. We also describe data we are unable to obtain, and the restriction this places on our model formulation and estimation.

Nielsen Videoscan collects DVD sales data from retailers at the point of sale. We use the weekly sales and price of all DVDs sold in the United States, aggregated nationally. Other researchers have used this dataset to study DVD sales (Elberse and Oberholzer-Gee, 2007). The dataset does not include Wal-Mart. In our period of interest, Wal-Mart was a major retailer of DVDs that carried a smaller inventory of possible titles than comparable national retailers. Hence, our sample may understate the importance of larger titles and overstate the importance of smaller titles. We supplement this dataset with estimates of print and advertising expenditure on movies at theatrical release from SNL Kagan. We lack data on print and advertising expenditure (P&A) by studios on DVD. Therefore, we use production cost, P&A in the theatrical channel and box office revenue that are likely to be closely correlated with DVD P&A.

We do not observe release costs in our dataset. Costs in the motion picture industry are comprised predominantly of the production costs of a movie and P&A. Production costs are borne upfront prior to release of a movie in the theatrical channel, and do not affect the release timing of the movie. P&A costs vary seasonally, over time and by firm, and thus affect the release timing of the movie. In our model, we assume that release costs may be incurred by a studio both as a fixed fee for in-store promotions, and through retailer margins. For instance, the fixed release costs of releasing titles on DVD include the cost of in-store promotions in the post-release weeks. In-store P&A might cost more in weeks of peak demand, when retailers face maximum demand for in store advertising and shelf space. We assume titles do not face distribution constraints; this assumption is clearly more appropriate for this market than for the theatrical release market. More importantly, for reasons of tractability, we assume that the retailer plays no strategic role. Another piece of missing data is that we do not observe the weekly release announcements of studios in our dataset. Our model approximates an MPNE with announcements, without data on the announcements of studios.

Similar to Luan and Sudhir (2007), we restrict our study to titles released in theatrical channels prior to release on DVD to reduce computational load. We drop older titles released prior on VHS and re-released on DVD, from our sample. Some titles with smaller revenues, either low production cost sequels or children’s titles, may be released direct-to-DVD and are dropped from our sample as we expect the dropped titles have a limited competitive effect on the release timing and pricing game.[20] From 2000 to 2002, the subset of data used currently for estimation, we observe the release of 512 titles with 5339 observations of price and quantity post release.

In our model, prices (and release dates) are chosen by firms, given seasonal demand and release costs, and their rivals' announced and actual release dates and prices. There is considerable price variation in DVDs that cannot be predicted from title characteristics; a regression of price against title characteristics has an adjusted r-squared of 0.2237 (Table 5).

Table 5: Price Regression[21]

	Estimate	Std. Error
Intercept	3.29E+01 ***	5.82E-01
Weeks Since Release (WR)	-3.29E-02 ***	9.09E-03
WR^2	5.61E-03 ***	1.64E-03
WR^3	-2.73E-04 **	8.45E-05
log(Box Office)	3.72E-01 ***	2.74E-02

---

[20] We can incorporate the effect of dropped titles in the model. However, the computational cost of additional data is overwhelming and the lack of observables on smaller titles makes demand estimates noisy.

[21] We suppress coefficients for movie characteristics, weekly fixed effects, distributor fixed effects.
Signifiance codes: **'***' 0.01 '**' 0.05 '*' 0.1**
Multiple R-Squared: 0.2263, Adjusted R-squared: 0.2237

The strategic role of price is an important distinction between DVD releases and theatrical releases. In movie theaters, the price of a ticket is fixed regardless of the popularity of the title (Einav and Orbach, 2007). Hence, the two opposing forces when setting theatrical release dates are the lure of a peak demand week and the competition expected in that week. In our paper we study the joint evolution of two strategic variables (controls), release dates and prices, set simultaneously. The trade-offs between two strategic choices leads to outcomes that may appear to be anomalies when considering either variable independently. For instance, consider two movies: a blockbuster and a small independent movie (indie). The blockbuster is released in a week of peak demand, and the indie on a week of lower demand. Intuitively, we might expect the blockbuster to be priced higher than the indie. However, we find a negative correlation (-0.15) between total DVD sales in a week and the average release price of a new movie in the week. Thus the indie may be released at a higher price than the blockbuster. The joint modeling of strategic decisions allows for an explanation. In weeks of lower demand, there is lower short run competition. Hence, titles released in these weeks have higher release prices while those released in higher demand weeks, have lower equilibrium release prices.

Two empirical facts simplify our analysis and estimation. First, the (retail) price of a DVD at release is maintained over the first few months after release, with no significant decrease after release. This is significantly different than previous findings for prices of video games and other entertainment media, where the prices of titles after release decrease over time (Nair, 2007)[22]. For instance, regressing log price against time after release and other explanatory covariates, finds that the price of a title decreases by 6% on average over the course of the first 12 weeks (Table 1).

The second industry feature is that prices for DVDs are well approximated by discrete levels, allowing us to treat price as a discrete variable rather than a continuous variable. In appendix 2, we discuss relaxing this assumption and treating price as a continuous variable. Before we begin a formal discussion of the model, we discuss the timeline of firm actions.

3.2: Timeline of firm actions

In our model, each title released in the theater is a potential entrant in the DVD channel. In each week, a potential entrant may choose to either announce a price and week of release of a title, defer the announcement or change its previous announcement, including withdrawing the announcement altogether. Titles update their decisions simultaneously every week. The state of the industry is described by announced release dates and prices and actual release dates and prices. Pre-order forms from the Video Software Dealers Association indicate that final release dates and prices for DVDs are circulated to video stores, 4 weeks prior to the release of the DVD. Hence, we assume that the final release date and price decision is taken 4 weeks

---

[22] There are two explanations for the uniform price of a DVD for the first months, post release. Decreasing prices may lead to forward looking behavior from customers, who may wait for a price decrease and not purchase the DVD at the time of release. Store price guarantees, typical of retailers of home entertainment media, might make it unprofitable for a retailer to decrease prices after release.

prior to the observed final releases in our implementation but suppress the period in our notation.

A limitation of our study is that we specify a model at the level of a title, and ignore portfolio optimization concerns of a studio: cannibalization, the effect on release costs from multiple releases and the effect of strategic decisions on other formats and channels (such as the theatrical channel and rentals). Studios managing multiple titles may choose to spread DVD release dates to mitigate the effect of cannibalization and substitution, and/or choose to cluster DVD release dates to lower release costs. We measure payoff changes with earlier/later release, independent of revenue on DVD, to capture the net effect of the release decision over all channels. However, we cannot disentangle between the sources of the payoff variation: substitutability with the theatrical channel, changes in future revenue streams, etc. While our framework and estimation methodology allows for these issues, the additional computational burden is overwhelming in our application. Last, we ignore the role of the retailer, and model the studio as the profit-maximizing agent responsible for release strategy choices.

In the model, the value of a choice of an announcement implicitly includes the strategic value (either cheap talk or serious signaling) of making announcements, and accounts for both cooperative and competitive incentives. Titles maximize profits by choosing optimal announcements of release date and prices, in the presence of seasonally varying payoffs, leading to time-varying best responses for any title (as described in the previous section). For instance, a title may be more likely to release a movie if future periods have decreased demand, than if future periods have increased demand; the seasonality of demand leads to seasonality in the set of potential entrants (see Figure 2). In our model, incumbents face no strategic decisions. That is, once a title enters the DVD channel, it becomes part of the absorbing state of the Markov process in the release timing game. In our dataset, a DVD on average collects 75% of revenue in the first 20 weeks post release with post-DVD release. In this period, the price remains remarkably steady, decreasing by less than 10% of the release price (see Table 1). Hence, it is sensible to model only release price setting.

Inter-release perishability implies that a studio only considers a finite number of periods after theatrical release for the DVD release of the movie.[23] Our model does not assume that all titles must be released on DVD and is general enough to identify titles released in theaters which cannot be profitably released on DVD. Thus, while we assume that all titles play the release timing and pricing game, we allow titles to choose to not release on DVD.

There is an important different between timing models and geographic competition models (see Seim (2007) and Vitorino (2007)). Both "classes" of models are interested in separately identifying the effect of (inter temporal and/or geographic) differences in profitability, and the role of competitors. However in a geographic competition game, two agents in a time period, either do or do not have a competitive effect on each other. This effect does not depend on when they entered, and is solely a function of the identity of each agent. In a timing game, competition is asymmetric inter-temporally. Movies released early do not face competition for the first weeks post release, from movies to be released later. Movies released later, face competition from movies

---

[23] Inter-release perishability implies that despite seasonal demand variations for any cost vector it is never profitable to release an unreleased movie after a finite number of periods. See assumption (A7) and lemma (L2) for a more complete treatment.

released early in the first weeks post release. The level of the competition faced diminishes with the gap between the release dates: older movies have a limited effect on newer movies.

To summarize: in our model, studios evaluate the value of release announcements in terms of resulting release schedules. The model nests a degenerate case of cheap talk where announcements communicate no information between studios. The costs and benefits of announcing are the changes in the industry landscape due to coordinating and competitive responses of other studios. We identify trade-offs between higher demand, competition, release costs and perishability by the choice of a studio to release the title on DVD, concluding the release timing and pricing game.

Our model specification is applicable when agents (firms) adjust dynamic decisions to changing industry landscapes. While unforeseen shocks, are accounted for in dynamic models that consider forward looking behavior (including our model), systematic industry changes of the nature described lead to a non-stationary MPNE. Other examples of such predictable shocks include market expansion, new product diffusion, changes in public policy, and release of complementor products. In the next section we describe the model specification.

3.3: Studio payoffs

We estimate the profit function per title per week. Profit is estimated in the expected two parts- revenue and cost. To identify the costs of releasing a DVD, we need to separate between the positive effect of pricing on profits from the negative effect of price on quantity sold. Reduced form profit functions based solely on the release timing schedule do not allow us to separate the profit into these components. We use a market outcome model to separate the effect of underlying seasonal shocks and competition on demand, from the seasonality of release costs.

In developing the sales (market outcome) model, we have two choices. We can model consumer demand from first principles of utility, accounting for dynamics in consumer demand (as did Luan and Sudhir, 2007). Alternatively, we can use a reduced-form capture of demand. We choose the latter for the following reasons. First, the focus of this paper is dynamics on the supply side. Researchers in this area typically use reduced-form models of revenue to simplify estimation (Bajari, Benkard and Levin, 2007). A model of firm market share in a period, allows for a parsimonious mechanism to account for the effect of competitors on per period profits. Second, our specification captures the relevant dynamics of inter-release and within-channel perishability, which are the two key dynamic elements that studios consider when setting release timing and pricing.[24] Finally, our data are aggregate, not individual-level (unlike Luan and Sudhir, 2007), making it less suited for structural demand estimation.

In appendix 1, we specify our non-stationary MPNE framework, and describe assumptions on model primitives and the resulting equilibrium. Our operationalization of the general frame work is presented below. Let p_{dwt}, x_{dwt} be the price and characteristics vector of DVD d, released in week w in time t.[25] To allow for

---

[24] We are unable to account for cross-channel substitutability; that is outside the scope of this paper.
[25] For convenience we index time as number of weeks since the first week of January 2000.

competitive effects while ensuring computationally tractability,[26] we model the market share of DVD d in week w and year y, ms_{dwt} as:

$$ms_{dwt} = \frac{\exp(\delta_{dwt})}{\exp(\delta_{dwt}) + \sum_{i \in \mathbb{C}_t \setminus d} \exp(\delta_{iwt})} \tag{6}$$

$$\text{where } \delta_{dwt} = \log(p_{dwt})\alpha + \log(x_{dwt})\beta + \xi_{dwt} \tag{7}$$

The market share model allows us to present a richer description of the industry. In estimating dynamic models, the effect of other agents is approximated by a linear function. In practice, the assumption either leads to an exponential increase (due to an increased number of agents) in the number of estimated parameters. Or symmetry restrictions on the profit function: competition being determined by state and not by identity. The use of a market share (market outcome) model alleviates data and computational requirements. In our application, we use descriptors of titles (box office revenue, genre, rating, and studio/distributor identity) when calculating the asymmetric competitive effects of titles.

Profits from releasing a DVD accrue post release. We write profits to a studio in period t from releasing a dvd d in week w as

$$\pi_{dwt} = p_{dwt} q(x_{dwt}, p_{dwt}) - f_{dwt}(q(x_{dwt}, p_{dwt})) \tag{8}$$

---

[26] Market expansion (e.g. Einav (2007)) and/or a random coefficients version of the market share model improve predictive capabilities but increase computational burden. In general, a model that fits the light tail conditions described in the paper can be used instead, without affecting the proof of convergence.

where $f_{dwt}\left(q\left(x_{dwt}, p_{dwt}\right)\right)$ is the marginal cost for dvd d, released in week w in time t. In our empirical application, we interact the movie and studio characteristics in a linearly additive specification:

$$\pi_{dwt} = \left(\lambda p_d - \gamma_V x_{dwt}\right) ms\left(x_{dwt}, p_{dwt}\right) f_t\left(Q_w\right) \tag{9}$$

where $\left(\lambda p_d - \gamma_V x_{dwt}\right)$ is the net average studio margin for dvd d, and with quantity calculated using the market share model and total weekly sales, $f_t\left(Q_w\right)$.

Perishability of the movie impacts both title payoffs and the competitive impact of the title. We account for diminishing appeal when calculating both the payoffs for th studio and the competitive impact of the movie on other titles. Further, as our titles differ across periods, the model adjusts to the changing sets of titles released.

The majority of a DVD’s revenue is garnered in the first months after release. In this period, competition and seasonality are major determinants of sales. Sales into the future are affected by competition from other movies released in the same week, but not from newer releases coming into the market in later periods. Hence, we model residual sales in remaining periods post the first 12 weeks, as a function of the seasonality of the week of release, the competitive set of the week of release and the observables of the movie. Thus, total payoffs to a studio from a title come from the first 12 weeks of profitability and a residual value of the movie:

$$\pi_{dw} = \sum_{j=w}^{w+M} \beta^{j-w} \pi_{dwj} + \beta^{M} \kappa_d\left(w\right) - \gamma_F x_{dw} f_w\left(Q_w\right) + \upsilon_{dw} \tag{10}$$

where $\kappa_d\left(w\right)$ is the residual sales and $\gamma_F x_{dw} f_w\left(Q_w\right)$ the release costs for dvd d, in week w. The general model allows υ_{dw} to be correlated across movies and weeks υ_{dw} .[27]

Our model payoffs are firm-specific. Ericson and Pakes (1995) specify a payoff function that depends solely on the number of studios in a particular state, and not the identity of the studios in that state. They model the state space using a set of counting measures to index the number of studios in a particular state. However, titles differ vastly in appeal. In our model the state representation is much richer and allows for observable differences between titles. The identity of a title impacts not only the payoffs of the DVDs but also addresses the impact of the title on other DVDs available concurrently. The affect of the title on the profitability of other titles depends on the composition of the choice set in that week.

4. Model Estimation

4.1: Challenges in solving for the MPNE

As mentioned in the introduction, we face three challenges in solving for the MPNE in our model. First, we lack intermediate release announcements of studios, and only

---

[27] The general forms of most extant dynamic models do not admit contemporaneous correlation as contemporaneous correlation biases estimates of the transition function.

observe the equilibrium final release schedule, leading to an under-identified transition matrix. Identification of the transition matrix and the use of Nested Fixed Point approaches require knowledge of the current and future state vectors, and the actions taken by agents. To ensure identification of the transition matrix, extant dynamic models have considered research questions where the state and action vectors can either be observed or imputed. While in our model we cannot observe all states and actions due to data constraints, in many applications such data remains unobserved due to other institutional details. For instance, privacy laws may prevent a store from identifying prior behavior of customers, censoring information on past decisions and their current state. Thus, we generalize dynamic models to settings where the researcher is faced with the burden of estimating on a censored dataset.

Second, while we prove the existence of a non-stationary MPNE in our model, we cannot solve for the general form of the MPNE as the state transition matrix is under-identified in a non-stationary MPNE. In the Ericson-Pakes (1995) frame work, identification depends on inter-temporal decisions of studios following a stationary Markov process (for a discussion on identification, see Berry and Tamer, 2006). In this frame work, the best response of a studio depends only on the industry state, and the state transition matrix is identified by the responses of studios to different industry states. In appendix 1.2, we describe the assumptions required to identify the model when best responses of firms change over time, and the related change in convergence properties.

Third, multiple equilibriums in a dynamic model require the specification of an equilibrium arbitration process over future equilibriums. The transition kernel and value function is unique to a particular equilibrium. Hence in a dynamic model with multiple equilibriums, agents in a period hold beliefs over which equilibriums will be played in future periods, to form expectations of future payoffs from a strategy. Without a methodology to arbitrate between equilibriums, the expectation over value functions is poorly defined. In general, randomizing between candidate equilibriums is in advisable as it rules out all signaling mechanisms between firms, including those based on observed variables. For instance, firms may know to play a particular equilibrium in periods of peak demand, and a different equilibrium in periods of low demand. It is also computationally intractable to enumerate possible equilibriums in the model and hence solve for the expected value of an action to an agent.

The presence of multiple equilibriums played in the data leads to inconsistent two step estimation (Aguirregabiria and Mira, 2007; Bajari, Benkard and Levin, 2007) due to the lack of a unique reduced form. A prior approach is to assume that while the researcher is unaware of the equilibrium selection process, and despite knowledge of the potential presence of multiple mixed and pure MPNE in the model, a unique equilibrium is played out in the data. This assumption is strong enough to both rule out equilibrium selection and inconsistent estimates of transition kernels (for instance, see assumption (5A) and (5B) in Aguirregabiria and Mira, 2007). In our data such an assumption is overly restrictive as it implies that all potential entrants in the six years of the dataset play the same equilibrium, across different holiday seasons, and in fast growing markets. The non-homogenous Markov kernel and choice function described and estimated in this paper are flexible enough to allow for the presence of multiple MPNEs in the data.

In the next section, we discuss a partial information estimation approach robust to all three issues described.

4.2: Solution Concepts

We draw from the solution concept of Oblivious Equilibrium (OE). In OE, agents are “oblivious” to the state distribution in a period, and optimize using Perfect Bayesian beliefs over candidate states. In the literature OE has been proposed in three separate contexts. First, in a model with a continuum of agents, MPNE and OE have been shown to be equivalent (Chakrabarty, 2003). Second, Krusell and Smith (1998) described a related model where agent behavior is derived from a response to the distribution of aggregate wealth rather than the precise allocation of wealth across agents, defending the solution concept as a behavioral model of agents in large markets. Third, Weintraub, Benkard, and Van Roy (2007) show that OE approximate MPNE models. They present error bounds for a model with a homogenous transition matrix and show that estimates of an OE converge to estimates from a MPNE in the context of large competitive industries where market shares decrease with the number of firms, particularly in a model using a logit market share function. In appendix 1.2 we show that given our choice of a logit market share function, estimates in our model converge to MPNE estimates despite the non-stationarity of the Markov kernel. We validate the model estimation by comparing agent actions forecasted with observed behavior, and compute an upper bound on the difference between OE and MPNE predictions.

Specifically, we replace the current state of the industry with a distribution over candidate state vectors that reflect the probability of observing the candidate vector in the time period. Rewrite the state vector as

$$\delta^t = \left\{ \delta_R^t, \delta_{UR}^t \right\} \tag{11}$$

where δ_R^t is the state vector for all movies released by time t, and δ_{UR}^t is the state vector for all titles unreleased at time t. Thus, we replace δ_{UR}^{t}, the unobserved state variables, with the distribution over candidate states consistent with Perfect Bayesian equilibriums. Equilibrium beliefs are neither imposed nor recovered from the data due to the under identified transition matrix.

The non-homogenous first order transition matrix requires us to write a period-specific choice function. We integrate over next period choice functions using the current periods' transition matrix $\widetilde{\psi}_t\left(\delta' \mid \delta\right)$ for each candidate vector and over all possible candidate vectors for the current period.

$$V_t^M(x_{it}, s_t, a_t, \upsilon_t; \theta) = \pi_{it}(x_{it}, \delta_t, a_t, \upsilon_t; \theta) + \beta E_{\delta_t} E_{\delta_{t+1}|\delta_t} E_{\upsilon} V_{t+1}^M(x_{i(t+1)}, a_t, \bullet; \theta) \tag{12}$$

(12) specifies a time-varying choice function in the model, due to the time-varying transition matrix. A non-stationary Markov strategy in the model for the studio is a function $\sigma_{it}^M : \Delta \times \nu \to A$. A non-stationary Markov strategy profile in model, σ_t^M is a set of non-stationary Markov strategies in the model, for each studio, period t. In the model, the necessary and sufficient equilibrium conditions are

$$V_t^M\left(\delta; \sigma_t^M\right) \geq V_t^M\left(\delta; \sigma_{it}'^M, \sigma_{-it}\right), \forall i, \delta, t, \sigma_{it}'^M \in I, \Delta, T, \Sigma^M \tag{13}$$

We draw on a strategy similar to extant static models of entry (eg. Bresnahan and Reiss, 1990) that use necessary conditions common to all equilibriums. The precise difference in strategic behavior of agents in our model and in a specific MPNE cannot be found without solving for the MPNE. Our model uses week-specific distributions over candidate state vectors without separating between MPNE equilibriums being played in the date or specifying the probability of playing a given equilibrium.

In appendix 1.2, we discuss the difference between our model and a stationary Oblivious Equilibrium (OE). Weintraub et al (2007) derive a theorem that shows that payoffs estimated using OE converge to payoffs estimated in a stationary MPNE. We discuss how to extend their results to our setting and show that payoffs found in the partial information estimator, converge to payoffs found in a full information non-stationary MPNE.

4.3: Model Estimation

As stated before, the OE estimation methodology proposed by Weintraub, Benkard and Roy (2007) is not appropriate for our setting due to the censored state space, non-homogenous Markov transition matrix and agent asymmetries in profit function. We first discuss an alternative characterization of the best response function and the resulting equilibrium, and then the use of these conditions in our estimation procedure.

In appendix 1.3, theorem (T1) implies agent playing the best response (making the optimal release announcement) in a period equivalently ensures the choice of an optimal absorbing state (final release date and price), accounting for the competitor's responses over the course of play. Hence (T1) shows that per period best responses can be translated into across period conditions on the choice of absorbing states. Note that the equilibrium description does not assume that other agents do not respond to the out-of-equilibrium actions of an agent. The assumption that actions, conditional on the path of play, are optimal across periods is a result of per period best response strategies of agents.

The re-characterization is intuitive: the underlying purpose of release announcements in the timing game is to ensure a path to the optimal period and price of release. Formally, in maximizing payoffs in a period, an agent engages in play to ensure that the course of play leads to the maximum payoffs for the agent, across multiple periods. The variation in profits from different release dates and prices stems from the seasonality of underlying demand and release costs, the endogenous evolution of competition as a response to the seasonal demand, and the effect of perishability on title profits.

Identification in the model is driven by comparing payoffs from releasing the movie, an absorbing state in the Markov process, with continuing in the game. This identification strategy has parallels to the literature in single agent dynamic programming problems where the agent has to decide the optimal stopping time (Rust, 1987), in an environment where payoffs from stopping vary over time. The optimal strategy in our model either prescribes releasing or deferring the release of the movie in a week, by maintaining a future release date, or choosing to postpone the release. An observed release indicates the studio found it optimal to maintain or choose the week and price as its announced release date and price respectively. Backtracking from the end of the finite planning horizon and recursively defining the value function, we implicitly construct the continuation value of deferring release. Hence, our estimator compares computed best response stopping points in the model and the decisions of studios to estimate trade-offs between equilibrium forces.

Our estimation and identification strategy is different from extant methods. In extant models, agent actions in a period, conditional on the Markov density, are best responses at equilibrium. In our model, the under-identified Markov density cannot identify the precise best response of the agent in the period. Instead, similar to prior Dynamic Stochastic Discrete Choice models (for a summary, see Ackerberg, Benkard, Berry, Pakes, 2007) we use a “two-step” method to calculate profits from conjectured releases. We forecast the industry evolution of the market, by modeling the evolution of the sufficient statistic. Titles that have been released have no further strategic decisions associated to them and are absorbing states in the Markov chain. The presence of an absorbing state and the forecasts of future industry environments, allow us calculate optimal release dates and prices. Next, we use estimates from the first period to find optimal release dates and prices of studios and maximize the quasi-likelihood.

However, in our model, unlike extant optimal stopping time models, the stopping decision of a firm depends on actions of other agents. In our approach, an agent makes an optimal decision while accounting for the behavior of other agents under the oblivious assumption. Thus, agent behavior in our model may differ from agent behavior in a MPNE. The use of a distribution instead of the information of actual agent state and hence future behavior implies an increased uncertainty which is manifest in the model as the difference in equilibrium outcomes in the MPNE and our model. The long term variance of the forecast is the sum of the long term or average variance of the true forecast generated in the MPNE, the residual variance of the forecasting equation and the average variance of the difference between the partial information approach and a complete information MPNE specification. Thus, the average difference between our model and the MPNE is bound by the mean sample variance of forecast errors.

While we can bound the degree of imprecision introduced by censored information (when compared to full information predictions), we cannot characterize the loss of efficiency in our model over a full information model. Our first stage estimates may be inefficient as they do not use the structural elements of the model. And we estimate the model without conditions on equilibrium actions, specifying the appropriate release announcement strategy, in each period. (T1) does not imply that our approach is econometrically efficient.

As the best response in our model is a unique strictly dominant strategy, the found equilibrium is a pure strategy Nash equilibrium in dominant strategies. Hence, econometrically an important difference between extant solutions of the MPNE model and our approach is that our estimators are econometrically complete in the presence of multiple equilibriums (Tamer, 2003). In extant dynamic MPNE models, multiple equilibriums make the MPNE model, even when the complete state and action space is observed, incomplete econometrically. In contrast, our estimators are consistent for all equilibriums and can be used without identifying equilibriums.

The estimation method described in the next section, uses (T1) and is general enough for any game with accrual of payoffs in periods after the choice of the absorbing state. This is a natural assumption in a game of release timing, where payoffs accrue post entry, but may not be a valid assumption in other games.

4.4: Estimation Algorithm

Our estimation algorithm has 4 steps:

Step 1: Market share Estimation

Estimate the market share model to scale. Let $\tilde{s}_{dwt}$ be the market share of dvd d in time t, in quantities. Define the geometric mean of in group market shares as $\ln(s_t^g) = \frac{1}{N_t} \sum_{i \in C_t} \ln(\tilde{s}_{iwt})$. Then from (7):

$$\ln(\tilde{s}_{dwt}) - \ln(s_t^g) = \ln(\delta_{dwt}) - \frac{1}{N_t}\sum_{i\in C_t}\ln(\delta_{iwt}) + \xi_{dwy} - \frac{1}{N_t}\sum_{i\in C_t}\xi_{iwt} \qquad (14)$$

Coefficients of the market share and residual sales models in our application are estimated using Ordinary Least Squares[28]. Identification and regularity conditions of the market outcome function have been well established in the literature. Parameter estimates from the first step are consistent in release timing games, but may not be consistent in entry/exit games. For instance, Pakes, Ostrovsky and Berry (2004), first estimate outcome values from exit decisions and then impute them in the second step of the estimation routine. The selective exit of firms in an entry/exit model may lead to a selection bias in the market outcome equation (first step), if estimated separately.[29] In our release timing game, almost all titles released in theaters are released on DVD and hence are present in the first stage of estimation, ensuring consistency.

Step 2: Forecast the sufficient statistic

We forecast a sufficient statistic to describe the effect of other agents on an agent's profits from release. An infeasible estimator can use the iterated Markov kernel to compute the described profit values of an agent across periods, under the assumption of optimal play. The kernel is under-identified in our application. Hence, instead we

---

[28] Clustering errors by week and using White's correction for heteroskedasticity does not improve fits and/or predictions.

[29] Entry selection bias can be corrected by using a control function of consistent estimates of the timing decisions of firms.

form a reduced form forecast of the evolution of the industry and assess the optimality of actions of studios when faced with the evolution of the industry.

A consistent forecast of a sufficient statistic can be formed in our research problem by looking at the seasonality of demand in a finite set of future periods. Inter-release perishability allows us to assume the existence of a finite end of the game, beyond which release is no longer profitable (see Lemma 2 in appendix 1.1). In our empirical application, we set nine months as the end of the release game and assume that titles which were not released nine months after theatrical release, exit the release game. By implication, agent decisions involve seasonality over the planning horizon, and the current level of the sufficient statistic.

In our application, we use $s_{fs} \triangleq s_t = \sum_{i \in \mathbb{C}_t} \exp\left(\delta_{it}\right)$ and forecast $E\left[Es_t\left(\mathbb{C}_t\right) | \delta_{UR}^t\right]$. The summary statistic is a measure of the number and strength of competitors, but is independent of the identity of competitors. Figure 3 shows empirical validation of the chosen summary statistic. In periods of peak DVD sales, the summary statistic is higher for released movies, indicating that the best movies were released. In contrast in periods of non peak sales, the summary statistic is higher for non released movies, indicating that the best movies were retained by studios for later release, in coming weeks of higher demand. The evolution of the industry is regressed on the current industry state, seasonality and future entrant vector:

$$\log(s_{dt}) = \theta_{fs} \log\left(s_{d(t-1)}\right) + z_t \theta_z + \tau \qquad (15)$$

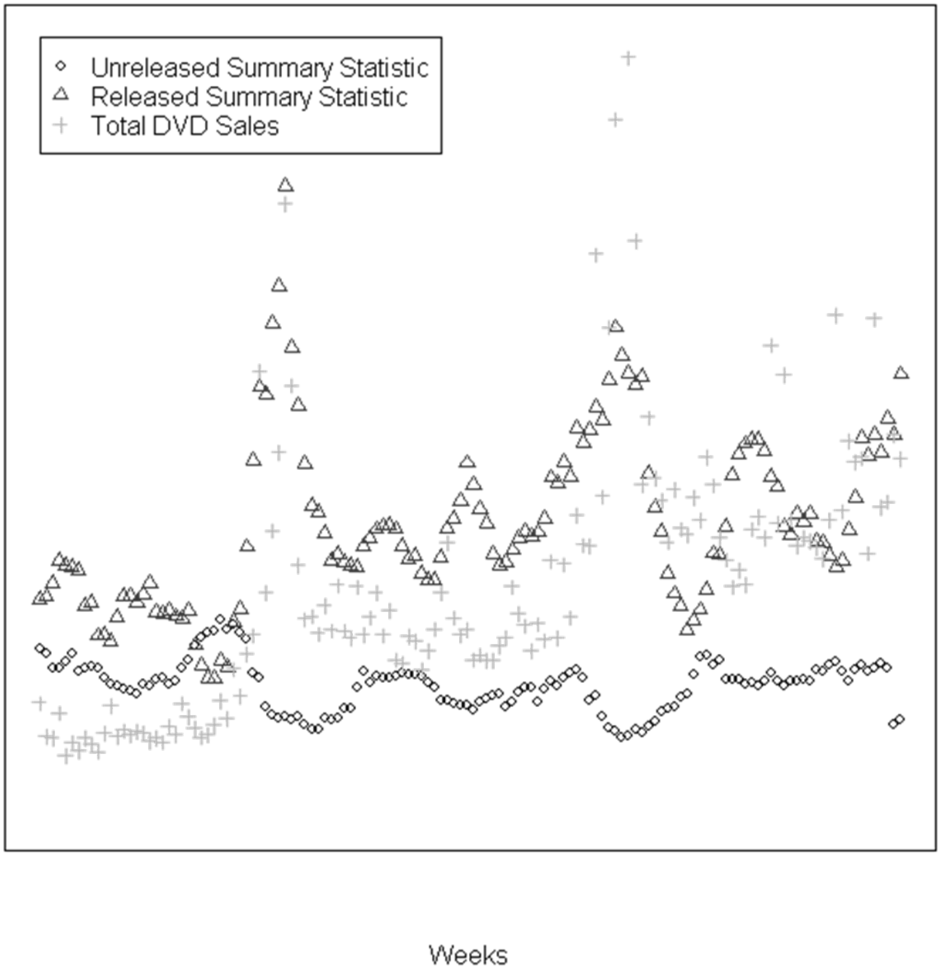

Figure 3: Total DVD Sales and Industry Evolution

Market share parameter (step 1) estimates are root-n consistent. Hence, forecasts of the sufficient statistic in our model are root-n consistent, and converge in probability to the true sufficient statistic.

Step 3: Compute sales from release dates and prices

We construct the empirical analog of the conjectured profits when releasing in a period. Forecasts from step 2, allow us to define expected payoffs from future actions. Using the market share model and a forecast of sales, we can compute expected total quantities of products sold for every given choice of release date and price. Hence for each agent, in every time period that the agent was in the timing game, we compute sales for feasible release date, price combination for the agent within the planning horizon.

Step 4: Maximize the Quasi-Likelihood

Our first stage estimates of the summary statistic are consistent, but in a finite sample are (with probability 1) not true parameter values. Using the sales estimated from step 3, we specify a quasi-likelihood estimation approach using a parametric specification of the payoff shock. Regularity conditions and other assumptions for the estimation are discussed in appendix 1.4. We use Richardson simplification to find $A(\theta_{ss})$ and the Eicker-Huber-White estimator for $B(\theta_{ss})$. If imputations of the summary statistic s_{fs} are heteroskedastic or autocorrelated, then standard errors of the sandwich estimator can be corrected by appropriately weighting the estimation function (Zeilies, 2006).We choose to use a quasi-likelihood-based method to maximize efficiency and ensure consistency of the standard error estimates. In an under identified model, similar to Bajari, Benkard and Levin (2007) one can instead follow Chernuzhov, Hong and Tamer (2007). Their estimator uses set identification to find parameters that describe difference equilibriums supported by the data minimizing a criterion function that penalizes violations of the best response function. The likelihood based approach is more efficient in the point-identified model, and produces precise standard errors of estimated parameters. In general, finding equilibriums in dynamic game models is computationally demanding. Most MPNE solutions increase exponentially in computational complexity and cost, with the number of agents in the model. In contrast, we are able to estimate on sets of potential entrants (on the order of 40 potential entrants in a period) larger than prior work on release timings as our model increases linearly in computational load, with the number of agents. The derived quasi-likelihood in our application is globally concave with closed form derivatives, further reducing computational load.

4.5: Identification

While the general framework of the model admits under-identified models, our model specification is point-identified. The identification of release costs comes from the effect on release decisions, of inter-release time, seasonal industry demand, and revenue from the title post release. Comparing across titles that could achieve the same revenue, we can identify differences specific to the attributes of the title. Specifically, studio margins are a function of the movie's characteristics (including inter-release time and time since release). The quantity sold in our model from a release date and price combination, is the product of the market share and the seasonal size of the market $ms\left(x_{smwt}, p_{smwt}\right) f_t\left(Q_w\right)$. A change in margins, affects profits depending on the revenue from the title. Hence the coefficients of studio margin are identified through the change in revenue with different choices of release strategies. γ_F is the vector of coefficients of the release cost function, identified through the change in industry sales of DVDs in the weeks post release. Studio differences in margins and release costs are identified in the model through differences in release behavior for similar titles in similar weeks, across studios.

The underlying variation identifying release costs are the different market shares and seasonal market sizes across different weeks, for different release dates and prices. The variation is induced by the underlying seasonality of demand, endogenous evolution of competition (choices of other studios) and perishability. There are two limitations to this approach: we cannot identify any release costs that are constant across the different weeks as they do not figure into the release timing optimization, and we are only identified to scale.[30] Our specification is similar to specifications used in complete information models. The described model frame work is general enough to allow variables unobserved by the econometrician but observed by agents (common un-observables). Common unobservables lead to decisions of agents being contemporaneous correlated. Similar to complete information models (Gallant, Hong and Khwaja, 2008), our estimators remain consistent under the assumption that common unobservables are orthogonal to observables but potentially correlated with private information shocks. However unlike extant complete information models we maintain restrictions on unobservables and shocks being independent over time. A complete information model assumes away elements of pre-emption and learning. The presence of private information potentially correlated with the common un-observables, implies that allowing serial correlation may lead to "learning" in the game described. The resulting model is beyond the scope of our research. Further, we restrict out attention to identification and estimation of our model in this paper for the parametric form. In future research, we plan to show that our model is semi-parametrically identified, and can be estimated using an extension of the approach of Hong and Shum (2007).

4.6: Results

We find that market share is well predicted by the print and ad spending of a movie,

---

[30] One can identify scale through the assumption on a discount factor that is less than 1. Identification to scale is adequate to build the counterfactuals and simulations that form the major substantive contribution of the paper.

which is positively correlated with larger box office revenues. Larger number of weeks since theatrical release significantly decreases the attractiveness of the movie.[31] As in Lehmann and Weinberg (2000), we find that larger box office revenue and greater screen-weeks exposure predicts higher consumer utility. Similar to Luan and Sudhir (2007), we find that longer inter-release times between channels decreases consumer utility. (See Table 6)

Table 6: Coefficients of DVD Market Share[32]

	Estimate	Std. Error
(Intercept)	-4.060 **	1.347
Price	-0.199 ***	0.021
log(Price)	1.674 ***	0.342
Weeks Since Release (WR)	-0.754 ***	0.026
Inter-release (IR)	-0.160 ***	0.019
log(WR)	3.702 ***	0.077
log(IR)	3.724 ***	0.496
log(Box Office)	0.909 ***	0.018
WR*IR	-0.0008	0.0009

---

[31] We try different time specifications and do not see a difference in fit across different specifications, including higher order polynomials of time spent in channel.

[32] Signif. codes: 0 '***' 0.001 '**' 0.01 '*'0.05 '.' 0.1 ' ' 1
Multiple R-Squared: 0.6019, Adjusted R-squared: 0.5871
We suppress coefficients for movie characteristics, weekly fixed effects, distributor fixed effects.

We regress the residual sales (sum of revenue in week 13 to week 24 after DVD release) on industry and movie characteristics (see Table 7). While the DVD market share in the first weeks after release is negatively affected by better movies, residual DVD market share is positively affected by the release of better movies at the same time as the DVD. The measured complementarity of titles may arise due to better releases increasings store visits to DVD retailers, and hence the number of older titles sold.

Table 7: Coefficients of Residual DVD Sales[33]

	Estimate	Std. Error
(Intercept)	5.31E+00 **	1.79E+00
Industry Sales (week released)	9.39E-08 *	4.23E-08
Industry Sales (11 weeks after release)	-4.74E-08	4.79E-08
Industry Sales (Average over weeks 13-23)	1.30E-07 **	4.10E-08
Industry Competition (week released)	1.63E-07	3.07E-07
Industry Competition (11 weeks after release)	1.53E-06 **	4.62E-07

---

[33] We suppress coefficients for movie characteristics, weekly fixed effects, distributor fixed effects. Significance codes: **0 '***' 0.001 '**' 0.01 '*'0.05 '.' 0.1 ' ' 1**
Multiple R-Squared: 0.826, Adjusted R-squared: 0.809

Table 7 (continued)

Unreleased Industry Competition (week released)	1.59E-05 **	6.05E-06
Unreleased Industry Competition (11 weeks after release)	1.57E-05	1.22E-05
Price	-5.78E-02 *	2.64E-02
log(Price)	8.21E-01 *	3.66E-01
Inter-release (IR)	-1.64E-02	2.58E-02
log(IR)	-3.60E-01	7.07E-01
log(Box Office)	9.63E-01 ***	3.07E-02

Estimates of DVD release costs (structural supply parameters estimated in the second stage) are in Table 8.[34] Our results indicate that release costs are seasonal, and are higher in weeks of peak demand. We find that movies that performed better at the box office, controlling for the increased sales on DVD, face lower net DVD release costs. We also find that movies that spend a longer time between the channels, have a longer inter release period, sell fewer copies due to the decreased market potential and incur higher release costs. Hence, a model ignoring release costs would overcluster optimal release predictions in weeks of peak demand, as it would ignore changes in release

---

[34] Due to the computational burden results have been estimated on data for 200-2002 for now. We will shortly estimate the model on the entire 2000-2005 data.

costs. Finally, both marginal release costs and fixed release costs differ across studios, genres and ratings. Our simulation results corroborate our prior explanation of the pricing anomaly. Regressing optimal release prices suggested by the model on seasonal demand variation shows that the model predicts lower prices for movies in weeks of peak demand. For every standard deviation increase in industry demand for the week of release, the simulation suggests a decrease of 16 cents in DVD release price.

Table 8: Coefficients of Release Costs[35]

		Estimate	Std. Error
Margins	Seasonal variation	14.58 ***	9.02E-01
	lg(Box Office)	-2.65 ***	9.15E-02
	Inter-Release (IR)	23.37 ***	5.13E+00
	IR^2	-6.69 ***	1.04E+00
Residual Sums	Seasonal variation	0.11 ***	2.81E-02
	lg(Box Office)	-0.008 .	4.74E-03
	IR	1.90 ***	1.22E-01
	IR^2	0.09 ***	2.49E-02

---

[35] Signif. codes: 0 '***' 0.001 '**' 0.01 '*'0.05 '.' 0.1 ' ' 1
We suppress coefficients for movie characteristics, weekly fixed effects, distributor fixed effects.

Table 8 (continued)

Fixed Release Costs	Seasonal variation	66.71 **	2.32E+01
	lg(Box Office)	-14.93 ***	2.52E+00
	IR	-222 .	1.32E+02
	IR^2	42.92 *	2.11E+01
IR	Seasonal variation	26.68 ***	3.16E+00
	lg(Box Office)	-2.72 *	1.12E+00
IR^2	Seasonal variation	-4.29 ***	4.94E-01
	lg(Box Office)	0.39 *	1.80E-01

4.7: Model Fit and Validation

The estimators of industry evolution are fairly accurate, indicating a reasonable model fit. In both Figure 4 and Figure 5, residual variance and forecast error is limited. The R squared of the 10 week future forecast equation is 0.9267 and R squared of the 30 week future forecast equation is 0.9114.

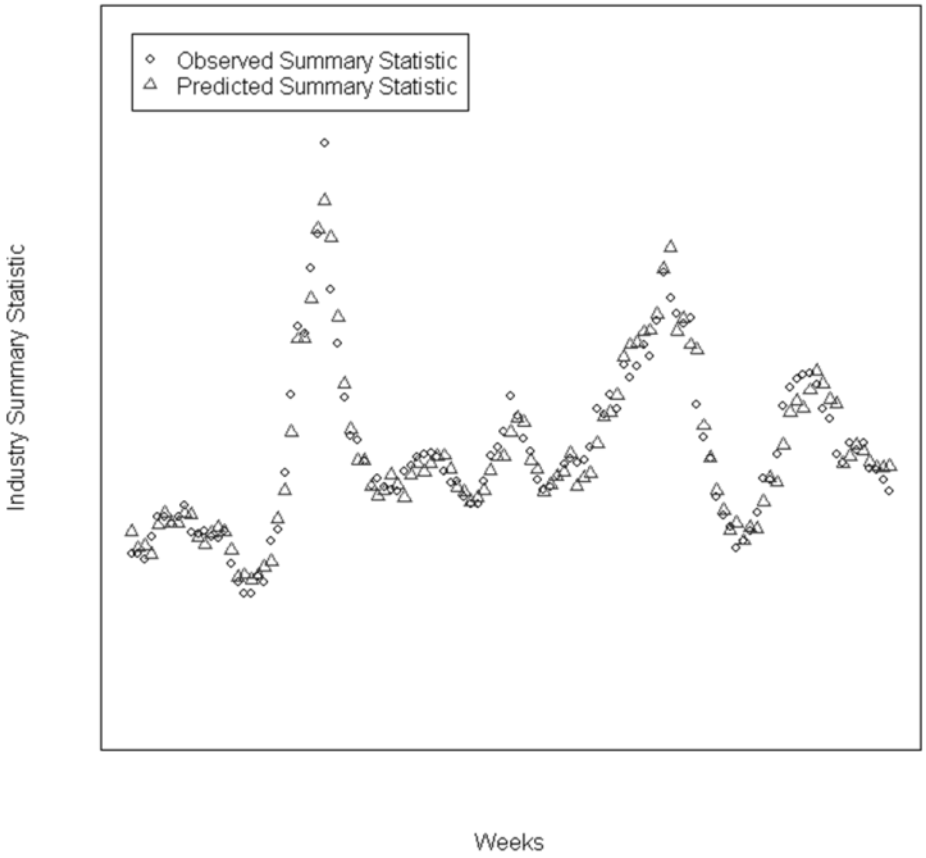

Figure 4: Forecasting 10 Weeks into the Future

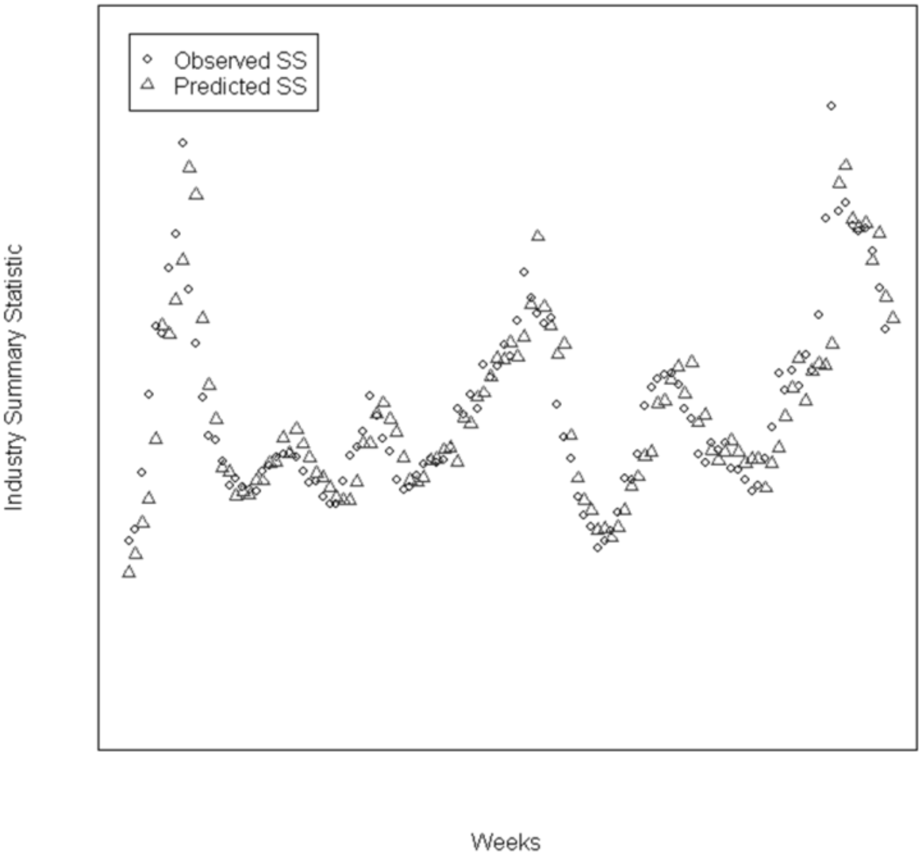

Figure 5: Forecasting 30 Weeks into the Future

We compared our in-sample model fits, with two alternative specifications:

i. M0: Reduced form model of prices and Theater-to-DVD window as a function of title characteristics

ii. M1: Dynamic model with release costs set to zero.

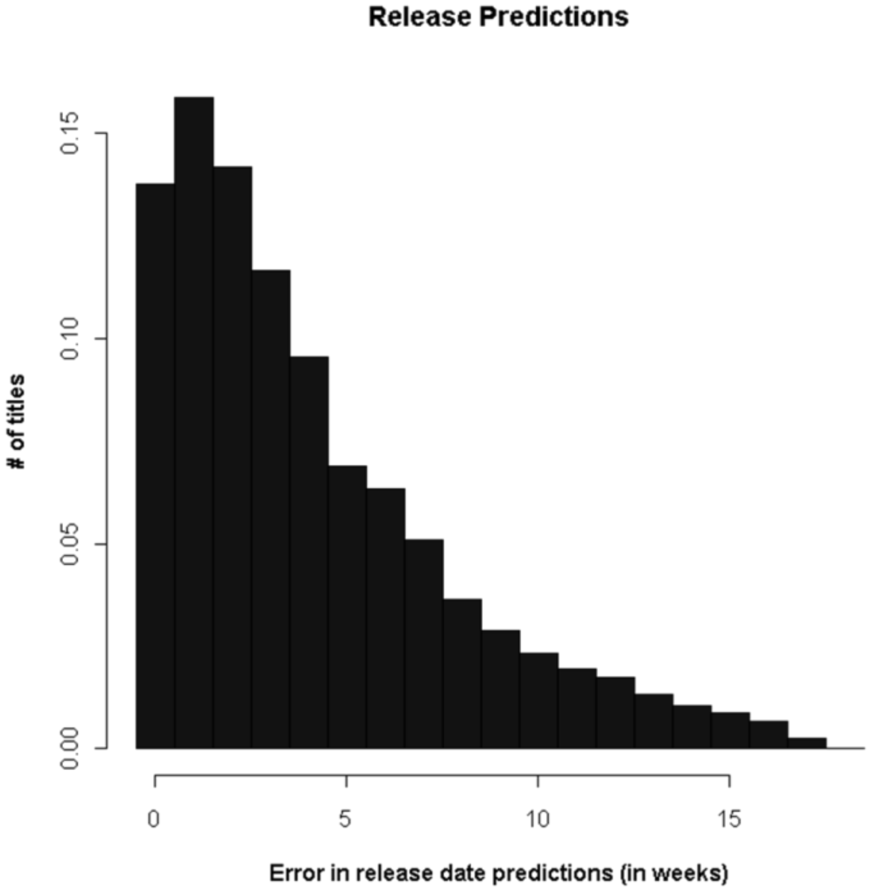

Figure 6: Histogram of Release Date Forecast Errors

The mean absolute error in predicting release dates for M0 is 4.73 weeks. Our model has a MAE of 4.05 weeks while M1 has a MAE of 4.27 weeks over the entire sample. Figure 6 is a histogram of absolute errors in release date prediction for our model, across the entire sample. For short term predictions (observed release in the coming 10 weeks), the model has a MAE of 2.86 weeks, while M1 has a MAE of 2.91 weeks. As expected the model performs better on nearer term than longer term predictions. We predict the release price with an accuracy of 52%[36].

5. Conclusion

DVD sales are a major source of studio profitability. As weekly sales vary dramatically over the year and the majority of sales for a title are made in the first weeks post release, the timing and pricing of a DVD release is a major strategic decision for studios. In this paper, we model the dynamic game of pre-emption and coordination played by studios when deciding the joint decisions of release date and price on DVD. In particular, we study the impact of seasonally varying demand, competition and release costs on the evolution of competition and release timing and pricing decisions, in the industry. An issue in estimation is that when setting release dates, studios use weekly announcements to mitigate competition in setting release and pricing schedules. Not accounting for these announcements might lead to biased estimates of the release timing and pricing game. We show how to account for these unobserved announcements to obtain robust estimates of this competitive timing and pricing games among DVD titles.

Substantively, we are able to measure unobserved release costs, allowing for both firm and title heterogeneity in the release cost function. We contribute to the literature

---

[36] We use 7 levels of price: \$0-\$5, \$5-\$10, …\$30 and above.

methodologically by developing estimation routines for models in which extant estimation methodology cannot solve for MPNE. We do not observe release announcements, and hence estimate the model on a censored state space. The policy functions of studios for determining release strategies change due to the variation in payoffs and the growth of the industry, leading to a non-stationary Markov process. Agent asymmetries prevent the use of counting measures for states to account for the impact of competition on studio profitability. Our estimators are econometrically complete, computationally tractable, and show reasonable predictive accuracy despite these constraints.

A limitation of our paper is that we assume that firms seek to maximize profits on DVD, ignoring positive network externalities on future channels and optimization over multiple titles. While theoretically the model scales to both multiple channels and portfolio optimization, a lack of data on other channels and the accompanying dramatic increase in computational cost, limit the empirical application.

A technical limitation of the model is that in using a simultaneous game of incomplete information, we are subject to the regret critique. Studios in our model make decisions on the basis of their own private information and beliefs on the actions of other agents, and cannot revisit their decisions. In contrast, in a sequential game, actions of rivals reveal private information, and hence may potentially lead to different best responses. These questions deserve further exploration in future research.

BIBLIOGRAPHY

Ackerberg, D., L. Benkard, S. Berry, and A. Pakes (2005), "Econometric Tools for Analyzing Market Outcomes," in *Handbook of Econometrics,* forthcoming

Aguirregabiria, V. and P. Mira (2007), "Sequential Estimation of Dynamic Discrete Games," *Econometrica*, 75 (1), 1-53

Ainslie, A., X. Dreze, and F. Zufryden (2005), "Modeling Movie Life Cycles and Market Share," *Marketing Science*, 24 (3), 508-17

Bajari, P., C. L. Benkard, and J. Levin (2007), "Estimating Dynamic Models of Imperfect Competition," *Econometrica*, 75 (5), 1331-70

Chiou, L. (2007), "The Timing of Movie Releases: Evidence from the Home Video Industry," *working paper,* Occidental College, CA, USA

Doraszelski, U. and A. Pakes (2006), "A Framework for Applied Dynamic Analysis in IO," *working paper*, Harvard University, MA, USA

Dorazselski, U. and M. Satterthwaite (2007), "Computable Markov-Perfect Industry Dynamics: Existence, Purification, and Multiplicity," *working paper,* Harvard University, MA, USA

Dutta, P. and R. Sundaram (1994), “The Equilibrium Existence Problem” in *General Markovian Games*, Cambridge: Cambridge University Press.

Einav, L. (2003), "Not All Rivals Look Alike: Estimating an Equilibrium Model of the Release Date Timing Game," *working paper*, Stanford University, CA, USA.

---- (2007), "Seasonality in the U.S. Motion Picture Industry," *RAND Journal of Economics*, 38(1), 127-45.

Elberse, A. and F. Oberholzer-Gee (2007), "Superstars and Underdogs: An Examination of the Long Tail Phenomenon in Video Sales," *working paper*, Harvard University, MA, USA.

Ellickson, P. and S. Mishra (2008), “Enriching Interactions: Incorporating Revenue and Cost Data into Static Discrete Games,” *working paper,* University of Rochester, NY, USA.

Ericson, R. and A. Pakes (1995), "Markov-Perfect Industry Dynamics: A Framework for Empirical Work," *Review of Economic Studies,* 62 (1), 53-82.

Foutz, N. Z. and V. Kadiyali, (2007), "Evolution of Preannouncements and Their Impact on New Product Release Timing: Evidence from the U.S. Motion Picture Industry," Johnson School Research Paper Series No. 11-07, available at SSRN: http://ssrn.com/abstract=1019531

Gallant, A. R, H. Hong, and A. Khwaja (2008), "Estimating Dynamic Games of Complete Information with an Application to the Generic Pharmaceutical Industry," *working paper,* Duke University, NC, USA.

Galloway, S. (2004), "Movies and the Media", *Hollywood Reporter*, retrieved from http://www.hollywoodreporter.com/hr/search/article_display.jsp?vnu_content_id=1000512483, on 10 September, 2008.

Glicksberg, I. L. (1952), "A Further Generalization of the Kakutani Fixed Point Theorem, with Application to Nash Equilibrium Points," *Proceedings of the American Mathematical Society*, 3 (1), 170-74.

Hong H. and M. Shum (2007), "A Semiparametric Estimator for Dynamic Optimization Models," *Economics Working Paper Archive 461*, The Johns Hopkins University, MD, USA.

Hennig-Thurau, T., V. Henning, H. Sattler, F. Eggers, and M. B. Houston (2007), "The Last Picture Show? Timing and Order of Movie Distribution Channels," *Journal of Marketing*, 71 (4), 63-83.

Krider, R. E. and C. B. Weinberg (1998), "Competitive Dynamics and the Introduction of New Products: The Motion Picture Timing Game," *Journal of Marketing*, 35 (1), 1-15.

Krusell, P. and Jr A. A. Smith (1998), "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 106 (5), 867-96.

Lehmann, D. R. and C. B. Weinberg (2000), "Sales through Sequential Distribution Channels: An Application to Movies and Videos," *Journal of Marketing*, 64 (3), 18-33.

Luan, J. and K Sudhir (2007), "Optimal Inter-Release Time Between Sequential Products: Application to Theatrical Movies and DVDs," *working paper*, Dartmouth College, NH, USA.

Mazzeo, M. (2002), "Product Choice and Oligopoly Market Structure," *RAND Journal of Economics*, 33 (2), 1-22.

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 (2), 129-53.

Pakes, A. and P. McGuire (2001), "Stochastic Algorithms, Symmetric Markov Perfect Equilibrium, and the 'curse' of Dimensionality," *Econometrica*, 69(5), 1261-81.

Prasad, A., B. Bronnenberg, and V. Mahajan (2004), "Product Entry Timing in Dual Distribution Channels: The Case of the Movie Industry," *Review of Marketing Science*, 2 (2004), Article 4.

Radas, S. and S. M. Shugan (1998), "Seasonal Marketing and Timing New Product Introductions," *Journal of Marketing Research*, 35 (3), 296-315.

Rust, J. (1987), "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55 (5), 999-1033.

Sawhney, M. S. and J. Eliashberg (1996), "A Parsimonious Model for Forecasting Gross Box-Office Revenues of Motion Pictures," *Marketing Science*, 15 (2), 113-31.

Seim, K. (2006), "An Empirical Model of Firm Entry with Endogenous Product-Type Choices," *RAND Journal of Economics*, 37(3).

Swami, S., J. Eliashberg, and C. B. Weinberg (1999), "SilverScreener: A Modeling Approach to Movie Screens Management," *Marketing Science*, 18 (3), 352-72.

Tamer, E. (2003) “Incomplete Simultaneous Discrete ResponseModel withMultiple Equilibria,” *Review of Economic Studies*, Vol 70, No 1, pp. 147-167.

Vitorino, M. (2008) , "Empirical Entry Games with Complementarities: An Application to the Shopping Center Industry", *working paper*, University of Chicago, IL, USA

Weintraub, G. Y., C. L. Benkard, and B. Van Roy (2007), “Markov Perfect Industry Dynamics With Many Firms”, available at SSRN: http://ssrn.com/abstract=870482

Zeileis, A. (2006), “Object-oriented Computation of Sandwich Estimators”, *Journal of Statistical Software*, 16(9), 1–16.

CHAPTER 3

MANAGING THE FORMAT TRANSITION FROM VHS TO DVD:

THE CASE OF PRICING AND AVAILABILITY OF VHS VERSUS DVD

1. Introduction

When industries find themselves at the cross-roads of technological changes, firms in the industry can make choices to speed along new technology diffusion or try to slow it down. In this paper, I examine the various incentives that firms face to support the old versus the new technology; these incentives stem from demand, cost and competitive dynamics. I address this question in the context of movie studios in the U.S. during 1997-2005 deciding whether to release movies or “titles” in the old platform of Video Home Disc (VHS), or the new Digital Video Disc (DVD) platform. DVDs provided better picture and sound quality than VHS tapes, and came with special features in addition to the primary programming content (movie). For simplicity and because of my current lack of data, I will abstract from the hardware or DVD player market in this paper. Therefore, the key drivers of consumer adoption and hence studio profits and technology diffusion in my model are whether studios released titles in VHS or DVD, and how they priced them.

To see how these firm choice variables influence their profits, for simplicity consider the incentive for any single firm in this industry, and consider two levels of prices-high and low- and two levels of product availability in the new format- high and low. If a studio releases several titles in DVD at low prices and its rivals continue to release in VHS, this studio will capture short-term profits from customer adoption. A disadvantage of this penetration pricing scheme is that in the longer term, consumers might not be willing to pay higher prices, and therefore a low entry price might set up self-fulfilling expectations of low prices in the future. An advantage of penetration pricing is of course a larger customer base. The studio also could have scaled back the availability of titles at this low price, but that would have slowed adoption.

Additionally, now consider competitive concerns a studio might face. The home video markets show indirect network externalities: customer adoption might have been faster if more studios had also released in this format. More studios releasing in this format also means direct cannibalization of demand. Similarly, if rival studios also adopt penetration pricing to encourage adoption, it is harder to coordinate a future price increase. Rival studios releasing many titles on a higher price might not help with adoption, or hurt this studio’s profits if consumers are willing to pay higher prices for rival (and even this studio) titles. Therefore, rivals’ actions can be simultaneously both strategic substitutes and strategic complements in this game with indirect network externalities.

Note too that there are asymmetries among studio incentives to push the new platform over the old one. First, larger studios with more titles might be better able to spread the cost of promoting the new platform over more titles than smaller studios. Also, some genres of movies (e.g. action movies) might benefit more from the new platform, whereas others do not (e.g. children’s movies due to the comparable indifference towards sound and video quality, and the simpler operation and robustness to rough handling of VHS tape).

To understand this complexity of competitive interactions, I employ a dynamic games framework to model studio pricing and platform choice for title releases. This is a natural tool to study my research question for several reasons. Firm's incentives when pricing and releasing titles depend on the current and future actions of competitors. Setting optimal prices and availability across multiple formats involves conjectures of present and future competition, adoption behavior (characterization of how many, and which consumers adopted), and industry support for the format. In my dynamic game, these variables are considered as states. Firms may either cooperate: set strategies to ensure consumer adoption, or choose to free ride on the efforts of other firms. Hence, the conduct of the industry determines pricing policies.

There are two challenges in estimating dynamic models with multiple equilibriums. First, extant dynamic models are econometrically incomplete when faced with multiple equilibriums, because the recovered estimate of the transition matrix is under-identified in the model. Second, agents in games with multiple equilibriums, hold beliefs on future equilibrium choices that define the value of each action. Past research has taken two approaches to this issue. First, researchers have assumed that the despite the potential presence of multiple equilibriums in the data, a single unique equilibrium is played in sample (Aguirregabiria and Mira, 2002). Second, researchers have constrained the strategy space in order to ensure the uniqueness of equilibriums estimated. Neither of these approaches is sensible for my application. In the presence of multiple possible equilibriums and especially in the presence of strategic substitutes and complements, outcomes observed will likely correspond to different equilibriums played out over the time period. Alternatively, to restrict strategic complementarity or substitutability requires demand function assumptions that are restrictive. This paper describes an admissible sub-game perfect selection mechanism that allows firms to arbitrate (choose) between equilibriums. The selection mechanism provides a (probabilistic) mapping between equilibriums: for any history of actions or states, the mechanism defines the probability of firms choosing a particular equilibrium. I show that if a particular equilibrium is chosen by other firms, it is sub-game perfect for the firm to choose the same equilibrium (the mechanism is sub-game perfect). Knowledge of the sub-game perfect mechanism in turn allows a firm to choose appropriately between actions in a period, using rational beliefs on future equilibrium choices. Hence, econometrically, the selection mechanism allows me to write the probability of an action in a period, both by restricting the beliefs of agents and weighting the policy functions in a period. Using the likelihood of actions, I describe two estimators: an efficient but computationally expensive full information maximum likelihood estimator and an inefficient but comparatively computationally inexpensive pseudo-likelihood estimator. Both estimators require Mathematical Programming with Equilibrium Constraints based optimization routines, which have been described in the recent literature (see Vitorino, 2008 for a summary). While my preliminary results support the estimation methodology, I limit the scope of this paper to a set of simulations to motivate the methodological contribution, and a description of the estimators. The simulations characterize the dependence of firm actions on key trade-offs between competitive and cooperative forces, in and across periods.

In future work, I intend to compare these estimators on both simulated and market data (described herein), to draw inferences on the dynamics of cooperative and competitive behavior in the industry. In particular the future goal is to look to answer: what were possible evolutionary paths for the industry? Which of these paths were chosen in the data and what governed these choices? What was the importance of cannibalization and substitution between formats, in determining the equilibrium actions? How would equilibrium pathways change if model primitives changed?

The equilibrium selection frame works, and suggested estimation techniques, are general enough to account for many other dynamic games with multiple equilibriums. My research question has direct parallels with the current format transition to Blue Ray (dominant high definition disc format) from DVDs. Generalizing further, several other markets display similar features e.g. digital cameras, memory, etc where not all firms need to use one compatible format. More broadly, the estimation methodology is useful for any industry where rivals both cooperate and compete.

The rest of the paper is organized hence. In §2 I discuss the related marketing and economics literature. §3 describes the data and §4 informally presents the game being modeled. §5 extends the MPNE model to allow for a multiplicity of equilibriums, and the econometric analysis of the estimators proposed. §6 discusses simulations of the model, and results obtained. The last section concludes.

2. Literature Review

I discuss extant literature in the next three subsections. First, I discuss related literature studying the movie industry. Next I discuss the literature studying technology platform battles in markets, and the strategic decisions of firms in platform markets. Last, I discuss the dynamic games literature and focus on my methodological contributions to the MPNE solution concept.

2.1: Movie industry

The movie industry has invited considerable attention from several marketing and economics scholars (see Eliashberg, Elberse and Leenders, 2006, for a summary). I focus my discussion here on models of competition within a exhibition channel (e.g. theatrical) and across channels (e.g. theatrical and DVD) in this industry

Three papers have proposed static models of within-channel competition, looking at the direct effect of release timing decisions of competitors on sales. Swami, Eliashberg, and Weinberg (1999) study multiplex screen allocation decisions and formulate a model to optimize exhibitor scheduling to mitigate the substitutive effects of competing movies. Ainslie, Dreze and Zufryden (2007), build a market share model that extends the BOXMOD model to study the lifecycle of a movie at the box office, measuring the lifecycle substitution effects of competition within a channel. Einav (2007) presents an empirical analysis of release timings in the U.S. movie industry, studying both seasonality and competition which incorporates the effect of seasonality in the study of competition. In a companion paper, these estimates are used to study the timing game, and optimal timings calculated for the industry (Einav, 2003). My paper studies release timing and pricing decisions in a dynamic context, where profits in future periods depend on current period prices and releases.

Competition across channels has typically been modeled for any single movie (see Lehmann and Weinberg (2000), Prasad, Bronnenberg, and Mahajan (2004) for theoretical models of such competition). For example, Luan and Sudhir (2007) model the impact of cannibalization of sales and rentals of movies, on box office revenues, accounting for forward looking behavior of the consumer at the theatre. My paper additionally studies strategic decisions for multiple titles across multiple channels, extending insights to settings with portfolio optimization. For instance, while it may be optimal for a single studio to deviate and shorten the theater-to-DVD window, as concluded in Prasad, Bronnenberg and Mahajan (2004), cheating on the industry compact may lead to other studios punishing the studio by releasing early in future games. Hence longer theater-to-DVD windows may be sub-game perfect equilibriums in a dynamic setting.

Two papers study movies in a multiple channel setting. First, Hennig-Thurau et al, 2007, use individual level discrete choice data to study the effect on studio profitability of different configurations of sequential distributional channels, optimizing release timings across these channels. As their goal is to study hypothetical configurations vastly different from current market conditions, they use conjoint data to model channel substitution, without accounting separately for either complementarities or market expansion. Second, Chiou (2007) models seasonal demand variation in secondary channels, controlling for competitive interactions within the rental revenue channel and in DVD and VHS sales. While both papers study the effect on demand of different product release strategies, my paper models the resulting supply side game due to the direct and indirect demand effects.

2.2: Platform competition

Diffusion models studying the adoption of platform products, or those with (direct and) indirect network externalities, have a rich history in marketing and offer an excellent way to capture competitive dynamics in a reduced-form way (for a review see Hauser, Tellis and Griffin, 2005). Other papers have explored consumer decision-making in such markets. For eg: Basu, Majumdar and Raj (2003) measure the externality sensitiveness of attributes, and hence the difference in externality imposed on different products. My paper differs from these by modeling the supply-side or competition among firms in a structural model and in more detail.

The paper closest to mine is Gupta, Jain and Sawhney (1999) who use a latent class probit model to model a consumer's adoption decision and then study the actions of competing firms in the High Definition Television market. Conceptually, they also seek to address the trade-offs between decisions supporting and hindering the adoption of the new platform. They address their question at a new product, prior to launch and use stated preferences (firm responses captured through a modified Delphi method) to model the evolution of the industry. My research question examines industry behavior post hoc to understand the role of competition, substitution and complementarity dynamics in the decisions of firms. My model is richer in being extendable to more strategies and in explicitly modeling the role of exogenous fluctuations in both consumer demand and supply, on equilibrium firm actions.

Other papers have proposed structural models of markets with platforms. In Nair, Chintagunta and Dube's (2004) model, consumer's adoption decision of a gaming console depends on the availability and prices of games (titles) in the current period. They specify a static supply-side game where they measure both the effect of software availability on hardware demand, and the effect of hardware demand on software availability. They assume a symmetric equilibrium, unlike my model where the differentiation among products is reflected in their optimal price and release strategy. Nair (2007) models the game between forward looking consumers and forward looking firms, in a monopolistically competitive industry (the video game software industry). He does not model demand complementarities, i.e. products do not contribute to the adoption of the format. This assumption is not a good one for my industry where titles can be strong substitutes in consumption and complements in the adoption of the technology. Dube, Hitsch and Chintagunta (2008) estimate their model on the video game industry, where hardware firms set the prices of consoles and hence determine adoption. Software firms are undifferentiated in their model, and provide software until the market is saturated. In contrast, the focus of my paper is the conflict in interests in a network-externality situation between short-term decisions to maximize adoption and long-term profitability of the platform in a market with multiple firms.

Hence my model addresses a gap in the extant structural literature in building a dynamic competitive model for portfolio optimizing strategic decision. I explicitly model the coexistence of strategic complementarity and substitutions in firm actions when firms play in both the old and new technology platform markets.

2.3: Dynamic games

Markov Perfect Nash Equilibrium models have been used to study marketing strategic decisions, which depend both on current and future choices of competitors. Most extant papers have focused on extending the Ericsson and Pakes (1995) frame work for analyzing dynamic entry and exit decisions taken by multiple firms in an oligopoly. In the original application, entry of firms was assumed to have a negative effect on other firms (strategic substitutes). Since then, their framework has been extended with methodological advances, to cases where the strategic space is much richer. A summary is in Dorazelski and Pakes (2007).

As mentioned previously, in models where strategies may be strategic substitutes or strategic complements, multiple equilibriums are a real possibility. The question of multiple equilibriums has remained unresolved for two reasons. First, extant dynamic models are econometrically incomplete when faced with multiple equilibriums: the transition matrix is under-identified (Tamer, 2003). The likelihood of an action requires accounting for choices in all equilibriums, and the likelihood of all equilibriums, and cannot be formulated in the extant framework. Second, agents in games with multiple equilibriums, hold additional beliefs on future equilibrium choices that define the value of each action. To ensure consistency, past approaches to multiplicity of equilibriums have either assumed that the despite the potential presence of multiple equilibriums in the data, a single unique equilibrium is played in sample (Aguirregabiria and Mira, 2007), or constrained the strategy space to ensure uniqueness (Ericsson and Pakes, 1995). Neither method of ensuring consistency is well-suited to my empirical context. In the presence of multiple possible equilibriums, one anticipates in datasets that stretch over the course of 6 years, outcomes observed will correspond to different equilibriums, leading to beliefs of agents on future paths of play, similar to the beliefs on the development and evolution of the future states in equilibrium. Alternatively, to restrict strategic complementarity or substitutability requires restrictive demand function assumptions. Therefore, to account for multiple equilibriums, I define a new framework that includes an equilibrium selection mechanism, and define consistent estimators that allow multiple equilribirums to be played in the data.

3. Data

In this section I discuss the data that motivates the paper, and will be used in future work to estimate the presented model. As mentioned earlier, the essay develops a frame work broad enough to capture the competitive dynamics of pricing and availability in the VHS to DVD migration. While the paper proposes a novel estimation strategy, and preliminary estimation results are encouraging, computational constraints limited its scope to a set of simulations the show the effect of key trade-offs discussed in the paper.

My data comprises release dates, quantities and prices of titles per week after release, as well as title-specific descriptors (e.g. box office revenue, etc.) for all VHS tapes and DVDs released in the United States between 2000 and 2005. I describe below the sources of these data and issues with them. I also describe data I are unable to obtain, and the restriction this places on my model formulation and estimation.

Nielsen Videoscan collects VHS and DVD sales data from retailers at the point of sale. I use the weekly sales and price of all VHS and DVD sold in the United States, aggregated nationally. Other researchers have used this dataset to study VHS and DVD sales (Elberse and Oberholzer-Gee, 2007). The dataset does not include Wal-Mart. In my period of interest, Wal-Mart was a major retailer of VHS and DVDs that carried a smaller inventory of possible titles than comparable national retailers. Hence, my sample may understate the importance of larger titles and overstate the importance of smaller titles. I supplement this dataset with estimates of print and advertising expenditure on movies at theatrical release from SNL Kagan. I lack data on print and advertising expenditure (P&A) by studios on VHS/DVD. Therefore, I use production cost, P&A in the theatrical channel and box office revenue that are likely to be closely correlated with VHS/DVD P&A.

I do not observe release costs in my dataset. Costs in the motion picture industry are comprised predominantly of the production costs of a movie and P&A. Production costs are borne upfront prior to release of a movie in the theatrical channel, and do not affect the release timing of the movie. In my model, I assume that release costs may be incurred by a studio both as a fixed fee for in-store promotions, and through retailer margins. For instance, the fixed release costs of releasing titles on VHS/DVD include the cost of in-store promotions in the post-release weeks. I assume titles do not face distribution constraints; this assumption is clearly more appropriate for this market than for the theatrical release market. More importantly, for reasons of tractability, I assume that the retailer plays no strategic role.

Similar to Luan and Sudhir (2007), I restrict my study to titles released in theatrical channels prior to release on VHS/DVD to reduce computational load. I don't consider older titles released prior on VHS and re-released on DVD. Some titles with smaller revenues, either low production cost sequels or children's titles, may be released direct-to-VHS/DVD and are also not considered in the simulations.

4. An Informal Outline of the Model

4.1: Description of the dynamic model

Studios are the agents in my model, and they choose whether or not to release a title on VHS/DVD and the price of the title if it is released. Each week, studios decide (simultaneously) on each title in their portfolio that has already been released in the theater, for potential entry in the VHS/DVD channel. Studios maximize profits by choosing optimal release dates and prices. As mentioned prior, studios managing multiple titles may choose to delay VHS/DVD releases to mitigate the effect of cannibalization and substitution, and/or choose to release more VHS/DVDs to lower costs. However, the model does not disentangle between all sources of the payoff variation: substitutability with the theatrical channel, changes in future revenue streams, etc. While the framework and estimation methodology could allow for many of these issues through the use of a richer consumer demand model, the additional computational burden is overwhelming in my application. Additionally, I ignore the role of the retailer, and model the studio as the profit-maximizing agent responsible for release strategy choices.

Incumbents face no strategic decisions in my model. Once a title enters the VHS/DVD channel, it becomes part of the absorbing state of the Markov process in the release timing game. In my dataset, a VHS/DVD on average collects 75% of revenue in the first 20 weeks post release with post release. In this period, the price remains remarkably steady, decreasing by less than 10% of the release price. Hence, it is sensible to model only release price setting.

4.2: Payoffs and Incentives

In the model, players choose actions variables in each period, contingent on the observables that affect their profits and the state of the world. To reduce the state space and make the model realistic, the evolution of future states is then additionally allowed to depend on the actions of consumers and realizations of exogenous shocks unforeseen by agents. I formalize this intuition by specifying two components of the model:

4.2.1: Consumer Adoption

The consumer adoption function tracks the adoption of DVDs by consumers as a function of studio strategies. For instance, the current price of DVD titles sets consumer expectations of future prices. I draw from the generalized Bass model to write adoption in time t as a function of the covariates (including firm actions) x(t), cumulative adoption until t, and the total potential market size, M:[37]

$$a_t = \alpha_0 x_t + \alpha_1 x_t \sum_{k=0}^{t-1} a_k + \alpha_2 x_t \left(M - \sum_{k=0}^{t-1} a_k \right)^2 + \xi_a \tag{13}$$

---

[37] I use a(t) for DVD hardware player sales, s(t) for DVD sales. x(t) includes state variables including price.

4.2.2: Studio Profits

State variables in a dynamic model summarize the effect of past strategic actions of firms, on current period profits and choices. That is, the state variable in the model is the link between choices across periods. In the model I use proxy for the effect of past release and pricing decisions on the market, in defining the current number of players and the mean price level.

$$Adoption\, index = \#of\ DVD\ players * \frac{mean\ price\, on\, DVD}{mean\ price\, on VHS}$$

I utilize a sufficient statistic for the adoption model:

(A1) Effects of actions are described by an industry summary statistic set (s_{fs}). There exists a consistent estimator $\hat{\mu}_{s_{fs}} : \hat{\mu}_{s_{fs}} \xrightarrow{as} \mu_{s_{fs}}$, where $\mu_{s_{fs}}$ is the true distribution of the summary statistic in a future period.

Assumption (A1) is similar to assumptions made in Bajari, Benkard and Levin (2007). Instead of assuming a finite parameter vector in the first stage of estimation, I assume the forecasted adoption rate from the first-stage converge to the rational beliefs of agents.[38]

I specify a parametric specification of the payoff shock, and similar to Bajari, Benkard and Levin (2007), to reduce computational load assume:

---

[38] Additional rate of convergence and local smoothness assumptions are required if using a criterion function for estimation as in Bajari, Benkard and Levin (2007).

(A2) The profit function, conditional on the demand function defined, is linear in unknown parameters. $\pi_{it}\left(x_{it}, p_{it}, \delta_t, a_t; \theta \mid \mu_{s_{fs}}\right) = \psi_i\left(x_{it}, p_{it}, \delta_t, a_t \mid \mu_{s_{fs}}\right) \bullet \theta$, where $\psi_i(\bullet)$ is a finite dimension vector of "basis functions" (including polynomial and interaction terms).

Assumption (A2) allows me to approximate the payoff function locally. A violation of (A2) does not prevent estimation or affect identification of the model. The described estimation methodology is robust to the use of a non-linear specification. As observed in Bajari, Benkard and Levin (2007), having a payoff function that is linear in unknown parameters implies that the constructed value functions are also linear in unknown parameters, simplifying estimation.

4.3: Mixing between equilibriums

Dynamic empirical game theory frameworks are useful as they provide a generalizing theory of oligopoly behavior (Fisher, 1989). The folk theorem shows that static demand and supply do not construct all equilibrium paths, as threats governing maximum deterrents (punishments) can support choices unavailable in static equilibrium. For instance, in a repeated prisoner's dilemma, strategies can be sustained which would not be sustainable in single episodal (static) game. In marketing, the substantive interest in dynamic games stems from the ability of dynamic game theory models to simulate the evolution of an industry. When a large number of equilibriums are often possible in these dynamic games then the recovery of the selection criteria is essential to forecast, and to test counterfactuals.

Why might firms mix between equilibriums? Different equilibrium outcomes lead different payoffs (profits). Hence, firms most likely would favor one equilibrium outcome over another. For instance, the largest player in the game may ensure that its most profitable equilibrium is played (Jia, 2008). Changing common unobservables, such as capital market pressures, may lead to changes in industry behavior. Alternatively, firms may try to equitably shape the industry profits, and ensure that they do better than the worst equilibrium for each.

Example 1:

Example 1 demonstrates the intuition behind profit sharing between equilibriums. Suppose players are 1 and 2. Player 1 chooses between {R,L}, and player 2 chooses between {A, B}. Table 9 shows the payoff matrix. Let {r,l} represent the probability of 1 playing R and L respectively, and {a,b} represent the probability of 2 playing A and B respectively. The two Pure Strategies Nash Equilibriums (PSNE) are {{1,0},{1,0}} and {{0,1},{0,1}}. The Mixed Strategies Nash Equilibrium (MSNE) is {{3/5, 2/5}, {1/7, 6/7}}. Write payoffs at the equilibrium as $\{\pi_1, \pi_2\}$ where π_1 is the payoff to 1 and π_2 is the payoff to 2. Equilibrium payoffs are in Table 10 below.

Table 9: Payoffs for Example 1

		2	
		A	B
1	R	30,10	0,0
	L	0,0	5, 15

Table 10: Equilibrium Payoffs in Example 1

{{1,0},{1,0}}	{30,10}
{{0,1},{0,1}}	{5, 15}
{{3/5, 2/5}, {1/7, 6/7}}.	{30/7, 6}

The MSNE in this example has profits lower than each of the PSNE. However, it is unclear as to which PSNE will be chosen. {{1,0},{1,0}} favors 1, while {{0,1},{0,1}} favors 2. Hence, suppose firms decide to make equal profits by coordinating on {{1,0},{1,0}} with probability p, and {{0,1},{0,1}} with probability 1-p. Solving for equal profits, I find that p = 2/7. The equilibrium selection of equitable profit sharing makes economic sense because the profits at equitable sharing are 2/7*30 + 5/7*5 = 170/14, higher than 1's profits in the MSNE and in {{0,1},{0,1}}. And 2 is better off than in the MSNE and in {{1,0},{1,0}}.

The researcher cannot observe which equilibrium is being played in a period. Then the data will suggest that players played {{2/7, 5/7}, {2/7, 5/7}}, as a MSNE[39] in the model. However, If 1 plays R with probability 2/7 then 2 is strictly better off choosing B. If 2 plays A with probability 2/7 then 1 is strictly better off choosing R. Hence, {{2/7, 5/7}, {2/7, 5/7}} is not a MSNE in the game.[40] Other mechanisms for selecting equilibriums may include:

a. Randomizing (conditional on state or action vectors): This mechanism might allows firms to share profits.
b. Pareto-dominance, risk dominance, payoff dominance: Firms facing risk loving, risk averse, and risk neutral capital markets would prefer pareto-dominant, risk

---

[39] The ability to separate between equilibriums in the computed example is an artifact of mixing between two PSNE in which players assign all probability mass to different actions in different equilibriums.

[40] Mixing between MSNE and PSNE can be similarly confounded in the data as a MSNE.

dominant, and payoff dominant equilibriums respectively. Choosing accordingly between equilibriums allows firms to play in a period, the equilibrium most favored by their lenders and shareholders.

c. Equilibrium favoring dominant player (Jia, 2008): In this mechanism, market power is used to ensure that the chosen equilibrium is most favorable to the most powerful player.

In essay 3, I model the selection mechanism to compare the policy function in equilibrium with actions taken in the data. In essay 2, identification was driven by optimality across multiple periods, not in a period, using the solution concept of an Oblivious Equilibrium. The presence of multiple equilibriums in a dataset implies the transition kernel is time-varying: equilibrium choice determines state transitions. The non-homogenous transition kernel defined (§5.1, A7) is agnostic on the source of the time in-homogeneity and hence broad enough to allow for the effect of multiple equilibriums on state transitions.

When identifying off current period actions, the model is incomplete when I do not explicitly model the equilibrium selection mechanism. The extended form of the game corresponds to the actual decisions taken by firms. Current MPNE models imply the decisions require taking expectations over rational beliefs due to a unique equilibrium. Different equilibriums lead to different beliefs. Hence the agent must also resolve the uncertainty over which equilibriums will be played in the future to understand the future implications of an action. If a unique long run equilibrium is played in the data then the specification of the stationary MPNE (with a time homogenous transition kernel) is complete without specifying additional beliefs on equilibrium arbitration. However if firms play different equilibriums in different periods, then the extant model is incomplete. The remainder of this paper develops an extension of the extant

MPNE model that allows multiple equilibriums to be played in the data.

5. Model

In this section I first define a model that assumes the existence of a unique equilibrium. This model is similar to the model discussed in essay 2. In §5.2 I discuss signaling, and in §5.3 extend the MPNE model to a multiple equilibrium setting. §5.3 describes problems with extant estimation methods, while §5.4 and §5.5 describe the two novel estimation approaches proposed. §5.7 concludes with a discussion on identification.

5.1: Unique Equilibrium MPNE Model

I assume model primitives are common knowledge to potential entrants and incumbents:

$$\left\{\pi_i(x_{it},\delta,a,:),\Psi(\delta'\,|\,\delta,a,:),\nu_{it},\beta\right\}_{(i,x_{it},\delta,a)\in I\times\aleph\times\Delta\times A}$$

Following prior empirical work (Doraszelski and Pakes, 2007), I restrict my attention to symmetric and anonymous equilibriums. A set of functions, is symmetric if $f_i\left(x_i,x_{-i},\delta_i,\delta_{-i}\right)=f_j\left(x_i,x_{-i}\delta_i,\delta_{-i}\right),\forall i,j$. Hence I abstract from the identity of the agent in the payoff function, and write $f\left(\bullet\right)$. The function, $f\left(\bullet\right)$ is anonymous if $f\left(x_i,x_{-i},\delta_i,\delta_{-i}\right)=f\left(x_i,x_{perm(-i)},\delta_i,\delta_{perm(-i)}\right)$, where $perm(-i)$ is any permutation of the indices of other studios. Note that symmetry and anonymity restrictions do not assume that studios are identical, but instead that studio and product differences are observed. Reduced-form game payoffs for all agents at equilibrium are a function of its characteristics, state vector and competitive set.

The state space is $I \times \aleph \times \Delta$ where the set of agents is $I \in \mathbb{Z}_+$, $\aleph$ is the Cartesian product of observed title characteristics and $\Delta \subset \mathbb{Z}_+^{\infty}$ is the set of agent states. $\pi_i(x_{it}, \delta_t, a_t, :)$ is the profitability of agent i when it has characteristics $x_{it} \in \aleph$, δ_t is the state vector at time t and a_t is chosen by agents[41]. $\Psi_t(\delta' \mid \delta, a, :)$ is the transition function that determines state transitions. I adopt the convention of using primes to denote subsequent period variables, e.g. δ' to denote δ_{t+1}, δ'' to denote δ_{t+2}. In an abuse of notation, I also use agent subscripts to denote the partition of the state and action vector describing an agent, e.g. δ_i to denote the state of agent i.

Incomplete information models simplify the analysis of the equilibrium (Seim, 2007), and are more likely to accurately represent the industry (e.g. given the non-standard contracts for sharing revenue and for deciding the promotional expenditure). Hence, similar to entry models, I assume that prior to making a decision studios receive a vector of private payoff shocks ν_{it}, drawn independently over time, from a distribution $G_\nu(\bullet \mid \delta_t, x_{it})$ with support on $\mathbb{R}^{|E_t|}$, and ν_t the collection of private shocks for all titles in period t. Private information shocks describe payoffs variations from strategic decisions, including changes in the costs of advertising, promotional expenditure and the manufacturing cost for the DVD/VHS. Finally I specify a discount factor β.

I impose further restrictions on the model primitives.

---

[41] My notation for action space is consistent with Molinari et al (2008) but inconsistent with some extant models. I refer to the action space as $A = A_1 \times ... \times A_I$ whereas other papers may reserve the term action, for the specific action space of a particular player, A_i.

(A3) The state space is finite ($I < \infty; x < \infty, \forall x \in \aleph; T < \infty; \delta < \infty, \forall \delta \in \Delta$).

(A4) Profits are bounded ($\underline{\pi} < \sum \pi_{it}(\bullet) < \overline{\pi}$).

(A5) Studios discount future payoffs $\beta \in (0,1)$.

(A6) Private information appears additively in profit function $\pi_{it}(x_{it}, p_{it}, \delta_t, a_t, v_t, :) = \pi_{it}(x_{it}, p_{it}, \delta_t, a_t, :) + v_{it}$. F_v is distributed absolutely continuous to the Lebesgue measure.

(A7) State transition, follows a non-stationary[42] first order Markov process with time in-homogenous transition function $\Psi_t(\delta' | \delta, a, :)$. In general, future states are a time varying stochastic function of past states and actions $\mu_{st}(\delta') = f_t(\delta, a_t), s.t. \forall \left\{\delta' \in \Delta, \delta \in \Delta^C\right\} \exists a$ with $\mu_{st}(\delta') = f_t(\delta, a) > 0$.

Assumption (A3) stipulates the finiteness of the state space. First, as I restrict my attention to titles released in movie theaters and ignore direct-to-DVD sales, my set of agents is always finite. Second, in practice, characteristics of a title have finite range. Third, perishability implies that in any time period, a firm only considers a finite number of future periods for release. Correspondingly, I restrict the decision vector and action vector of potential entrants to be finite. Assumptions (A4), (A5) and (A6) are features of commonly-used profit functions in empirical studies. The second part

---

[42] A time in-homogenous Markov process does not have a stationary long term distribution of states and actions. I use the phrase non-stationary MPNE and time inhomogenous MPNE interchangeably to refer to the model outlined.

of (A3) is appropriate in a game of release timing where the agent can only make profits post release of the title, or sale of the player.

Assumption (A5) describes the stochastic monotonicity and continuity requirements on payoffs, fundamental to the existence of MPNE. Relaxing distributional assumptions on private information increases the number of mixed MPNE supported in the model. Assumption (A7) requires the state of the world to evolve in a first order Markov process. The evolution of the next period's states, conditional on the actions and states of agents in the current period, is stochastic. In keeping with extant papers, I require states in future periods to be accessible from continuation states. (A6) and (A7) are implicitly equivalent to requiring additive separability and conditional independence of controls and errors. Finally my state space is finite (unlike Ericsson and Pakes (1995), who study a problem with an infinite state space), and hence, I do not require agent payoffs to be bounded at the extremums of the state space. Assumptions (A3 – A7) lead to the following lemma:

Best Response Lemma (L1): Generically, the best response function is a unique mapping from $I \times \Delta \to A$, referred hereafter as $BR : I \times \Delta \to A$.

Proof: (L1) follows naturally under assumptions (A3), (A4), (A5), (A6), (A7) and the solution concept of a Perfect Bayesian Equilibrium (PBE). In a PBE, a studio considers expected profits from each strategic choice. The expectation on the profit function includes probabilities on the decisions of incumbents and potential entrants in future periods, described by the transition kernel specified in (A7). When taking expectations, indifference between two actions, making the best response function a correspondence occurs on a set of measure zero due to the continuity restrictions imposed in (A6).

--

Corollary (CL1):

If a firm is in a continuation state, $\delta_i \in \Delta_i^{C}$, then all future states of the firm have positive probability. $\int\limits_{a_t} \mu_{st}\left(\delta' \mid a,\delta\right)\mu_a\left(a\right) > 0$

Proof: (A6) implies that from each state, all actions are played with positive probability. (A7) implies that generically, all future states are achievable from continuation states.

--

For a particular equilibrium, (L1) implies a unique mapping from a combination of the state vector and observables to a future state. If agent actions lead to stochastic state changes, then agent beliefs in the PBE are rational, as (L1) implies (A7).

A stationary Markov model assumes that best responses depend only on the current state of the agents. Doraszelski and Satterthwaite (2007), prove the existence of a MPNE, and under certain conditions, the existence of a Pure Strategies MPNE. Assuming a stationary MPNE, they write the choice value function as

$$V(x_t, a_t, \delta_t, \upsilon; \theta) = \pi(x_t, a_t, \upsilon; \theta) + \beta E_{\delta_{t+1}|\{\delta_t, a_t\}} E_{\upsilon} V(x_t, a_t, \delta_{t+1}, \upsilon; \theta) \qquad (14)$$

where expectations on future value functions are taken over possible next period states, using the transition matrix.

A stationary Markov strategy for a studio is a function $\sigma_i : \Delta \times \nu \to A$. A stationary Markov strategy profile σ is a set of stationary Markov strategies for each studio in a period. The necessary and sufficient equilibrium conditions in a stationary MPNE are

$$V\left(\delta;\sigma\right) \geq V\left(\delta;\hat{\sigma}_i,\sigma_{-i}\right), \forall i,\delta,\hat{\sigma}_i \in I,\Delta,\Sigma \tag{15}$$

A non-homogenous first order transition matrix requires me to rewrite the choice value function. I write (15) as a period-specific choice value function, taking expectations over the next period choice value functions using the current periods' transition matrix

$$V_t(x_t,a_t,\delta_t,\upsilon;\theta) = \pi_t(x_t,a_t,t,\upsilon;\theta) + \beta E_{\delta_{t+1}|\{\delta_t,a_t\}} E_\upsilon V_{t+1}(x_t,a_t,\delta_{t+1},\upsilon;\theta) \tag{16}$$

(15) specifies a time invariant choice value function, while (16) specifies a time-varying choice value function[43]. Time-varying choice value functions, particularly when lacking estimates of the transition matrix, cannot be analyzed using extant methods without arbitrary restrictions on V_t. As current decisions are affected by future seasonality, for instance to control for the effect of seasonality one would need to make V_t a function of future periods.

A non-stationary Markov strategy for a studio is a function $\sigma_{it}:\Delta\times\nu\rightarrow A$. A non-stationary Markov strategy profile σ_t is a set of non-stationary Markov strategies for

---

[43] Blackwell's theorem does not apply to the general class of non-stationary Markov Perfect Nash Equilibriums. For instance, consider an infinite period game in which the market grows faster than the discount rate. I assume an upper bound on the profit function and implicitly a starting condition where no DVDs have been sold. These two conditions used with backwards induction arguments, guarantee the existence of the choice value function.

each studio in period t. In a non-stationary MPNE, the necessary and sufficient equilibrium conditions are

$$V_t\left(\delta;\sigma_t\right) \geq V_t\left(\delta;\hat{\sigma}_{it},\sigma_{-it}\right), \forall i,\delta,t,\hat{\sigma}_{it} \in I,\Delta,T,\Sigma \tag{17}$$

Assumption (A8) formalizes the intuition behind the chosen modeling functions described in §4.2 being time-homogenous. Similar to Essay 2, I assume the integrated value function $E_{\delta'|\delta}V_t(x_{it},\delta_t,\upsilon_t;\theta) = E_{\delta'|\delta}V(x_t,\delta_t,\upsilon_t,\wp_V\left(t\right);\theta')$ and the Markov kernel, $\Psi_t(\delta' \mid \delta,a_t,:) = \Psi\left(\delta' \mid \wp_\psi\left(t\right),\delta,a_t,:\right)$, are stochastic functions of $\wp_\psi\left(t\right)$ and $\wp_V\left(t\right)$, finite cardinality function vectors. For a discussion on identification of time inhomogeneous (non-stationary) MPNE, see Appendix 1.

Time Homogeneity Assumption (A8): Firm profits are time homogenous conditional on state and independent variables.

(A8) requires that profits in a period can be modeled using the state of the industry (adoption rate, price elasticity or any other measure of demand) and exogenous variables (such as seasonality). In DVDs, studio profits in a period are a function of the titles released (and their prices), and the adoption of DVDs until the period, satisfying the assumption. The assumption would be violated if there exists a structural break in the dataset (for eg, if the adoption function changed form in the middle of the study).

5.2: Signaling

In general, an infinite number of selection mechanisms (probability distributions) are admissible in this problem. The selection mechanism model used empirically allows all previously outlined situations. My formulation is more general than that in the extant literature. For instance the equilibrium drift model (Binmore and Samuelson, 1999) cannot handle the alternating equilibrium model. Even relaxing their Lipschitz continuity requirement does not allow the model to track the “pure best response dynamics” of the industry, the goal of my approach.

This paper modifies the definition of admissibility presented in Beresteanu et al (2008). Aumann expectations are expectations over probability mappings, using the probability of the mapping being played in the data. In equilibrium the policy function is a probability vector that defines equilibrium actions, while the selection probability governs the probability of selecting the equilibrium. Hence the expected policy function is the Aumann expectations over the set of policy functions.

As in Beresteanu et al (2008), I define admissibility through the set of probabilities on observables generated by the mechanism. Intuitively I seek to allow (admit) mechanisms which generate the conditional probabilities of actions and states observed in the data. The definition assumes the selection mechanism is common knowledge, with agents having rational beliefs on how the future will evolve. The admissible mechanism weights likelihoods derived from policy functions (of different equilibriums) to match the probability of the data, while maintaining rational belief structures of agents. More formally:[44]

DEFINITION: An Admissible Selection Mechanism in the dynamic game satisfies:

---

[44] I refer the reader to Molinari et al (2008) for a more formal discussion and approach (using random closed sets).

a. The probability of observing an action conditional on the information set I_t at t is $\Pr\left(a \mid I_t\right) = \int_{\Omega} \sigma^e\left(a \mid x, \delta\right) \mu_\omega\left(e \mid I_t\right)$, the integral of the policy function $\sigma^e\left(a \mid x, \delta\right)$, over all equilibriums using the selection probability $\mu_\omega\left(e \mid I_t\right), \forall e \in E$. The calculated probability $\Pr\left(a \mid I_t\right)$ is equal to the observed conditional probability of actions, given the information set.

b. Beliefs of future equilibriums resulting from the selection mechanism are Perfect Bayesian.

To ensure an admissible mechanism can be defined in my model, I further assume:

Exogeneity Assumption (A9): There is an exogenous mapping between states and actions, unaffected by the equilibrium generating actions. $E\left[\mu_\delta\left(\delta'\right)\right] = E\left[\mu_\delta\left(\delta' \mid x, \delta, a\right)\right]$, implying that the unconditional distribution of future states can be modeled in the data.

I assume that all actions relevant to the evolution of the dynamic game are observed, and modeled as being endogenous to the game. Formally, (A9) states that after agents in the model have chosen their actions, the evolution of (states in) the model is stochastically well defined. In most applications of game theory, states are a history of past actions. If the state is a stochastic (or deterministic function) of past states and actions, then (A9) is trivially satisfied. If state changes are affected by current period exogenous observables, for eg. adoption being a function of the economic environment in the period, then (A9) requires that once agents have chosen their pricing and release strategies, states evolve stochastically as a function of choices and the exogenous observables.

Drawing from the correlated equilibrium literature (Aumann, 1974), I assume ω is a coordinating common information shock that aids equilibrium selection (ω is visible to all players at the beginning of the period, prior to actions being taken). The remainder of this section shows that the coordinating/selection equilibrium for ω is stable in the sense of a sub-perfect equilibrium (SPE), and allows us to construct equilibrium selection probabilities.

Consider $\omega \in \Omega$ such that $\mu_{\omega}\left(\omega' \mid I\right)$ is the conditional probability measure describing a probability density on the common information shock ω'. Assume that Ω is equipped with a partition of $\left\{\omega_e\right\}$ such that $\mu_{\omega}\left(\omega' \in \omega_e \mid I\right) =$ probability of equilibrium e, conditional on the information set $I = \left\{x, \delta, a, e\right\}$.

Signaling Lemma 1 (SL1): For the game to be in equilibrium, the equilibrium selection rule must be a mapping from ω to the set of possible equilibriums: on seeing ω agents play a particular equilibrium with probability 1.

Proof: Suppose no selection rule can exist: on seeing ω agents face uncertainty over the equilibrium to be played, with no equilibrium played with probability 1. Hence, more than one equilibrium is played with positive probability, conditional on ω, implying agents face an expected policy function, integrated over the probability of each equilibrium:

$$E\left[\sigma_{-i}\right] = \int_{E} \sigma_{-i}^{e} \Pr\left(e \mid \omega\right) \qquad (18)$$

(18) is the Aumann expectations of the policy function, using the conditional probability of each equilibrium, $\Pr\left(e \mid \omega\right)$. From (L1) I know that the agent has a unique best response to the expected policy function $BR\left(\int_{E} \sigma_{-i}^{e} \Pr\left(e \mid \omega\right)\right)$. However, as no selection rule exists, the expected policy function cannot be the policy function of any of the equilibriums. Hence, the response function of the agent violates the

necessary conditions (15) for the model to be in equilibrium. That is $\left\{\int_E \sigma_{-i}^e \Pr(e \mid \omega), BR\left(\int_E \sigma_{-i}^e \Pr(e \mid \omega)\right)\right\} \neq \sigma^e, \forall e \in E$.

Now suppose that on seeing ω agents play a particular equilibrium with probability 1, implying $\Pr(e \mid \omega) = 1$ for some e. Hence, $\int_{e' \in E} \sigma_{-i}^{e'} \Pr(e' \mid \omega) = \sigma_{-i}^e$ and $\left\{\int_E \sigma_{-i}^e \Pr(e \mid \omega), BR\left(\int_E \sigma_{-i}^e \Pr(e \mid \omega)\right)\right\} = \sigma^e$. Thus, on seeing the signal all competing agents play according to the policy function of a particular equilibrium. As the best response of the agent is to play the action described in the policy function, a deviation is not profitable, implying that the model is in equilibrium.

--

Signaling Lemma 2 (SL2): If ω is unobserved, then the signaling rule is under-identified.

Proof: Consider an augmented state vector $\{\delta, \omega\}$. Identifying the selection rule requires observing or imputing the equilibrium played in each period. By (CL1), I know that all states are reached with positive probability in all equilibriums, implying that I cannot impute the choice of equilibrium played in a period. As in equilibrium, all actions are played with positive probability I cannot impute the equilibrium played in a period from observed states and actions.

--

Instead assume:

Signaling Assumption (A10) The equilibrium selected in a period only depends on the state and action history until that period: $\mu_\omega(\omega' \mid I) = \mu_\omega(\omega' \mid x, \delta, a)$.

(A10) implies that in the model firms can signal intent through their actions, ensuring competitive and coordinated choices.

Signaling Lemma 3 (SL3): A mapping from state to state ($\Delta \times A \rightarrow \Delta$) is formed (implicitly through ω).

Proof: (SL1) and (L1) together imply that in each period there is an equilibrium mapping from $\Delta \times A \rightarrow A$, conditional on ω. The transition kernel implies (SL3).

--

Signaling Lemma 4 (SL4): The data can be represented by a single selection rule between equilibriums, in any dynamic game with a SPE.

Proof: (SL1) implies that at least one admissible selection rule exists in the data. Proof by contradiction: suppose two different selection rules (mappings) are required to model the game played in the data. For agents to take rational decisions there must be an arbitration rule between the mappings. Define a meta-rule using the chain rule of probabilities between the arbitration rule and selection rules. Then this is a single selection rule which represents the game played in the data, contradicting the initial assumption.

--

Signaling Lemma 5 (SL5): States and signals are jointly Markov.

Proof: (SL3) implies that states are Markovian, conditional on the equilibrium selection rule and ω. From (A10), I know that the conditional probability of ω depends only on the last period's actions and states. Hence I get $\Delta \times \Omega \times A \rightarrow \Delta \times \Omega$.

--

5.3: Multiple Equilibrium MPNE Model

This section first defines the complete MPNE model, and then characterizes existence of equilibriums in the model. Last, it explores the needed assumptions and frame work extension. As the selection mechanism expands the scope of the model, the complete model primitives are:

$$\left\langle \mu, \pi_i(x_{it}, \delta, a, :), \nu_{it}, \beta, \left\{ \Psi_e(\delta' \mid \delta, a, :), \right\}_{e \in \mathrm{E}} \right\rangle_{(i, x_{it}, \delta, a) \in I \times \aleph \times \Delta \times A}$$

The complete model specifies different transition kernels across equilibriums (due to different policy functions) and an arbitration process between equilibriums.

Existence Theorem (ET): Given assumptions (A1 – A10), the following hold:

i. Beliefs are rational (PBE).
ii. An MPNE exists in the model.
iii. Any equilibrium, conditional on the private information shock, is a PSNE with probability 1.
iv. Multiple equilibriums (with different policy functions), but with a unique selection rule, may be played in the data.

Proof:

i. (SL1) restricts agent beliefs to Aumann expectations over (different) policy functions, using the selection rule. Hence, agent beliefs are rational expectations over future outcomes.

ii. From (SL5), states and equilibriums are Markovian. The model hence fits conditions described in Essay 2 (with ω as the censored state variable). From Essay 2, a MPNE exists in the model but that the general transition kernel remains under-identified in the data.[45]

iii. (L1) implies that given a set of policy functions, the best response mapping is generically unique; with probability 1, conditional on the private information shock, the agent has strict preferences. Hence, conditional on the private information shock, equilibriums are PSNE with probability 1.

iv. The game may have multiple equilibriums. (SL4) implies the selection rule is unique.

--

Existence Lemma (EL1): The Markov chain has a unique recurrence class of states and equilibriums.

Proof: Define the augmented state vector $\{\delta, e\}$. The model then matches Pakes and McGuire (2001), who showed that the (augmented) Markov process describing the state vector has a unique recurrent class.

(EL1) implies that in the long run, when describing the future states through the PML estimator, I remain in the recurrent class of states and equilibriums. Pakes and

---

[45] The proof is based on Dorazelski and Pakes (2007), who discuss and show the existence of MPNE in an infinite horizon model with uncountable states.

McGuire (2001) show that with probability one, forward simulations of the model land in the recurrence class. (EL1) hence implies that I can forward simulate multiple periods of information, instead of a single period, when estimating the non-structural components of the PML estimator.

--

Existence Lemma (EL2):

i. Beliefs are Martingale and converge.
ii. The integrated value function is well defined and <u>not</u> a function of the equilibrium.

Proof:

i. The Bayes map between states is well defined in the model. As the industry evolves, with continued diffusion of DVD technology, firms update their beliefs using the Bayes map. As shown in Easley and Kiefer (1988), firm beliefs are hence Martingale and converge.
ii. (SL3), (SL4) and (SL5) imply the integrated value function is unique across equilibriums. Proof by construction: any chosen action implies a distribution on future states. Conditional on the action and current states, I obtain a distribution over future equilibrium choices. Conditional on future states and current actions, define the Aumann expectations over all future action choices. From (i) I know that the beliefs converge despite industry evolution. Hence there is an integrated value function which only depends on current period actions and states.

--

(EL2) allows us to write a likelihood function for each action, in which the integrated value function is unique across all equilibriums.[46] The PML estimator builds a non-structural analog of the unique integrated value function, thereby reducing the computational burden of the estimation process. Further details are discussed in §5.6.

If beliefs are Markovian then the equilibrium selection mechanism is Markovian. This is more stringent than only admissibility: there may be admissible mechanisms that are not Markovian. The rationale for restricting attention to Markovian belief structures is provided by (RET), which shows that any Perfect Bayesian Markov belief structure requires a Markovian selection mechanism.

Reverse Existence Theorem (RET): Suppose a MPNE model with multiple equilibriums played in the data. Then:

i. The selection rule is Markov.

ii. States and equilibriums are Markov.

Proof:

---

[46] States in my model follow a pseudo-finite Markov process. Pseudo-finite processes can be shown to converge to a Markov process (Rosenthal, 1992). However, the likelihood of an action remains mis-specified despite the convergence result, if multiplicity of equilibriums is ignored in an extant MPNE model. In specifying a selection mechanism, the paper defines the weights of the pseudo-finite Markov process and hence defines the likelihood of each observed action.

i. Suppose not. Then the beliefs of the future would either involve more than the current information set, and hence not be rational, contradicting the assumption of a MPNE, or require knowledge of past states and actions (not just current states and actions), contradicting the Markov assumption on the transition kernel.

ii. Suppose not. From (i) the selection mechanism must be Markov. (SL5) shows that this implies the states and equilibriums must be jointly Markov, contradicting the initial assumption.

--

Hence (RET) shows that restricting to the Markovian signaling mechanism is consistent with the idea of a MPNE. A MPNE both leads to and relies on the Markovian structure of signaling. As one can only have a MPNE if the selection rule is Markovian, modeling Markov admissible mechanisms is consistent in the model. A limitation of the model is to assume that equilibriums evolve through the action space. (SL2) shows the full model to be under identified, and hence (A10) allows for estimation, while retaining the flexibility of the model.

5.4: Inconsistency of extant approaches

Often the researcher can define a set of inequalities that can be used for estimation even if from the observable and a realization of the unobservable, more than one event are possible. Beresteanu et al (2008) presents the efficient estimator for the sets of parameters that can support the data: the smallest set of parameters which solve the difference equations implied by the model (through the inequalities). Other common static multiple equilibriums papers build a series of inequalities that do not rely on a selection mechanism.

A dynamic model cannot be estimated in this way because of the extended form incompleteness. How does incompleteness impact estimation if one ignores multiplicity? In each equilibrium, the policy function is a different probability vector. From the same observables and un-observables multiple equilibriums and hence multiple policy functions are possible. The selection mechanism defines an appropriate method to weight between policy functions, and hence builds a consistent likelihood function.

While the nature of bias generated in extant models is hard to classify as the incompleteness either implies that the model makes unrealistic assumptions or is econometrically mis-specified, it is instructional to study the source of the bias. In particular, two estimation strategies have been used in extant papers:

a. Some papers restrict strategies to be either strategic substitutes or strategic complements. This approach is often theoretically appealing when considering non portfolio decisions, as in the extant literature. In my application, the assumption is unrealistic and would lead to mis-specification bias.
b. Other papers assume a unique equilibrium is played in the data. Either these papers use a:
    i. 2 step estimator (PML estimator): in the first stage, the policy function, state transition functions are estimated from the data. The first stage estimates are inconsistent when mixing between equilibriums, as shown in Example 1, as the observations do not represent mixed equilibrium observations.
    ii. Nested Fixed Point approaches (NFXP): the policy function and transition kernel is built from structural elements of the model. When using NFXP, the researcher assumes multiple periods have the same policy function, implicitly assuming the selection probability remains constant over time. The MLE is

inconsistent as the likelihood function does not incorporate the multiple policy functions, and hence is mis-specified in the model.

5.5: Defining a maximum likelihood estimator

Drawing from Vitorino (2008), and using (EL2), the probability of seeing an action can be written as:

$$l\left(a_i \mid x,\delta,\omega,\mu_{s_{fs}}\right)=\frac{\exp\left(E_{\sigma_{-i}}\left[\left(\pi(a,\delta)+\beta E_{\delta'|a}V(a,\delta')\right)\mid x,\delta,\omega,\mu_{s_{fs}}\right]\right)}{\sum_{\dot{a}_i\in A_i}\exp\left(E_{\sigma_{-i}}\left[\left(\pi(\dot{a},\delta)+\beta E_{\delta'|\dot{a}}V(\dot{a},\delta')\right)\mid x,\delta,\omega,\mu_{s_{fs}}\right]\right)} \tag{19}$$

$$\textit{subject to}\,\sigma\left(a \mid x,\delta,\omega,\mu_{s_{fs}}\right)=\prod_{a_i\in a} l\left(a_i \mid x,\delta,\omega,\mu_{s_{fs}}\right)$$

I build the probability of seeing the data (actions) by integrating the likelihood of the data in a potential equilibrium over all potential equilibriums. Hence, (19) leads to the Full Information Maximum Likelihood (FIML) estimator:

$$\hat{\theta}_{FIML}=\arg\max_{\vartheta\in\Theta}\prod_{d}\int_{\Omega} l\left(d \mid x,\delta,\mu_{s_{fs}},\omega;\vartheta\right)\mu_{\omega}\left(\omega \mid I_t;\vartheta\right) \tag{20}$$

(20) integrates the likelihood over all values of the common information shock ω. As agents choose between (discrete) finite strategies, I know that the model can only support finite PSNE. Hence, (SL1) implies (20) can be re-written as:

$$\hat{\theta}_{FIML}=\arg\max_{\vartheta\in\Theta}\prod_{d}\sum_{e\in E} l\left(d \mid x,\delta,\mu_{s_{fs}},\omega\right)\mu_{\omega}\left(e \mid I_t\right) \tag{21}$$

I estimate the model by maximizing likelihoods. I model $\mu_{\omega}\left(e \mid I_t\right)$ with iid Gumbel errors leading to a finite mixture logit model. Computationally, the difficulty stems from finding an integrated value function consistent with all best response strategies. In FIML, the solution requires a nested fixed point as shown in Pakes and McGuire (2001), and is a generalization of the FIML estimator described in Vitorino (2008). The value of the nested fixed point is different for each conjecture of the parameter vector. While a non-linear optimization routine (such as Knitro) builds shadow functions across the binding equalities, reducing the computational burden, the routine is not practical in problems with large state spaces and asymmetric agents.

To use a likelihood estimation approach, additionally I assume the following regularity conditions:

(R1) $\theta \in \Theta$ is a compact subset of $\Re^{|\theta|}$ and true value $\theta^0 \in \operatorname{int}\Theta$.

(R2) The pseudo-likelihood function $\Upsilon(\bullet \mid \mu)$ is uniquely maximized at θ^0, and $\Upsilon(\bullet \mid \mu)$ is twice continuously differentiable in $\theta \in \Theta$ with probability 1.

(R1) and (R2) are common regularity conditions for pseudo-likelihood estimation, met by the iid Gumbel specification of payoff shocks in my application. Under (A8), the argmax of the pseudo-likelihood function is a consistent estimator of the structural parameters. From (A4), the likelihood function $\Upsilon(\bullet \mid \mu)$ is continuous, leading to $\Upsilon(\bullet \mid \hat{\mu}) \xrightarrow{p} \Upsilon(\bullet \mid \mu)$. Maximizing the pseudo-likelihood yields second stage structural parameters (θ) whose variance is the sandwich estimator $A(\theta)^{-1} B(\theta) A(\theta)^{-1}$, where $A(\theta)$ is the Hessian of the log pseudo-likelihood and $B(\theta)$ is the variance of the pseudo-score.

For completeness I discuss the endogeneity of observables and a method for correcting for the endogeneity bias. Formally, I assume:

(A11) $E_{\nu}(\nu_{it}x_{it}) = 0$

A violation of (A11) would bias the coefficients estimated due to endogeneity. To correct for endogeneity bias in my model, when recovering the estimates of the adoption function, follow instrumental variable estimation methods. In the supply side game, violation of (A11) implies that both individual agent errors and forecast errors are correlated with agent observables. Grouping the error terms leaves a single error term correlated with observables. To estimate the model, I define a set of moment conditions by matching the best response with calculated best responses and interact the conditions with instrumental variables. While this method corrects for any potential endogeneity bias, it is econometrically less efficient than maximizing the best response pseudo-likelihood. One can use Richardson simplification to find $A(\theta)$ and the Eicker-Huber-White estimator for $B(\theta)$. If imputations of the adoption, choice or are heteroskedastic or autocorrelated, then standard errors of the sandwich estimator can be corrected by appropriately weighting the estimation function (Zeilies, 2006).

5.6: Pseudo Maximum Likelihood (PML) estimator

As the FIML estimator is computationally very expensive, this paper follows a PML approach. My method is a generalization of PML estimator described in Aguirregabiria and Mira (2002), hereafter AM. The difference between the FIML and PML estimator is that the integrated value function found as a nested fixed point in the FIML estimator for each conjectured parameter guess, is found using recovered probabilities in the PML estimator. The assumption in the PML estimator is the same as in the FIML, but the observed probabilities forces the likelihood for non true values to optimize using the optimal observed distribution. At true parameter estimates, both PML and FIML coincide.

Allowing for multiple equilibriums implies that I cannot use the PML estimator defined in AM. In AM, they use the policy function recovered from the data in the first step, to build the pseudo likelihood function. As the policy function is different across equilibriums, in my problem, a mixed policy function exists in each period (through the selection mechanism). Instead the described PML is a hybrid estimation strategy that uses results from AM.

AM describes an alternate representation of the fixed points in probability space. This alternate representation of the equilibrium uses the equilibrium probabilities of actions (Representation Lemma, AM). Both best response mappings $\{\Lambda, \Psi\}$ lead to identical fixed points. In a particular equilibrium there is a unique policy function of agents; conditional on ω I have a unique mapping $\sigma\left(a \mid x, \delta, \omega, \mu_{s_{fs}}\right) \triangleq \sigma^{e}\left(a \mid x, \delta, \mu_{s_{fs}}\right)$. To ensure notational equivalence with AM, define the integrated value function of the alternative mapping as:

$$\begin{aligned}\tilde{V}_i^{e*}\left(a, x, \delta; P^{e*}\right) &= E_{\sigma^{e*}}\left[\pi\left(a, \bullet\right) + e^{e*}{}_i\left(a, x, \delta\right)\right] \\ &+\beta \sum_{\delta' \in \Delta'} \tilde{V}_i^{e*}\left(a, x, \delta'; P^{e*}\right) f\left(\delta' \mid a, x, \delta\right)\end{aligned} \tag{22}$$

where $e^{e*}{}_i\left(a_i, x, \delta\right)$ is the conditional expectations on the private information shock, if a_i is picked as an action in the next period. Define $e^{e*}{}_i\left(a_i, x, \delta\right) = E\left[\varepsilon_i\left(a_i\right) \mid a_i = \arg\max_{a \in A_i} V^{e}\left(a_i, x, \delta, \varepsilon_i\left(a_i\right)\right)\right]$. Then re-arrange (22) as

$$\begin{aligned}\tilde{V}_i^{e*}\left(a, x, \delta; P^{e*}\right) - \beta E_{\delta' \mid a, x, \delta} \tilde{V}_i^{e*}\left(a, x, \delta'; P^{e*}\right) \\ = E_{\sigma^{e*}}\left[\pi\left(a, \bullet\right) + e^{e*}{}_i\left(a_i, x, \delta\right)\right]\end{aligned} \tag{23}$$

The profit function is linear in basis vectors (due to assumption (A2)). As errors are logit, I get $e^{e*}{}_{i}\left(a_i, x, \delta\right) = Euler's\,constant - \ln\left(P_i^{e*}\left(a_i, x, \delta\right)\right)$. Stacking(23), I get $\|A\| \times \|\Delta\|$ equations for the value function. To find the value function across equations, pre-multiply with the probability of the equilibrium, to get:

$$\tilde{V}_i\left(a, x, \delta; P^{e*}\right) = \sum_{e \in E} \tilde{V}_i^{e*}\left(a, x, \delta; P^{e*}\right) \mu_{\omega}\left(e \mid a, x, \delta\right) \tag{24}$$

In FIML, the algorithm has to solve for (24) explicitly. Both terms in (24) are defined through the econometric structure: the policy function and the selection mechanism are identified in the problem. To establish a parametric estimator, consider the stacked up set of equalities, and notice that I can solve equalities for all equilibriums. Pre-multiplying and adding, I get a stacked set of equations[47]:

$$\left(I - \beta F_{a\delta}^{*}\right)\tilde{V}^{*}\left(a, x, \delta'; P^{e*}\right) = E_e E_{\sigma^{e*}}\left[\pi\left(a, \bullet\right) + e^{e*}\left(a_i, x, \delta\right)\right]$$

$$\Rightarrow \tilde{V}^{*}\left(a, x, \delta'; P^{e*}\right) = \left(I - \beta F_{a\delta}^{*}\right)^{-1} E_e E_{\sigma^{e*}}\left[\pi\left(a, \bullet\right)\right] + \left(I - \beta F^{*}\right)^{-1} E_e E_{\sigma^{e*}}\left[e^{e*}\left(a_i, x, \delta\right)\right]$$

As,

---

[47] $F_{a\delta}^{*}$ is the expected probability of actions a, and states δ : elements of $F_{a\delta}^{*}$ are bounded between 0 and 1. As the discount factor is strictly less than 1, $\left(I - \beta F_{a\delta}^{*}\right)$ has an inverse.

$$\left(I-\beta F^{*}_{a\delta}\right)^{-1} E_e E_{\sigma^{e*}}\left[\pi_i\left(x_{it},p_{it},\delta_t,a_t \mid \mu_{s_{fs}}\right)\right]$$

$$=\left(I-\beta F^{*}_{a\delta}\right)^{-1} E_e E_{\sigma^{e*}}\left[\psi_i\left(x_{it},p_{it},\delta_t,a_t \mid \mu_{s_{fs}}\right)\theta\right]$$

$$=\left(I-\beta F^{*}_{a\delta}\right)^{-1} E_e E_{\sigma^{e*}}\left[\psi_i\left(x_{it},p_{it},\delta_t,a_t \mid \mu_{s_{fs}}\right)\right]\theta$$

$$=\left(I-\beta F^{*}_{a\delta}\right)^{-1} F^{*}_{a}\psi_i\left(x_{it},p_{it},\delta_t,a_t \mid \mu_{s_{fs}}\right)\theta$$

$$\Rightarrow \tilde{V}^{*}\left(a,x,\delta';P^{e*}\right)=\mathrm{M}\theta+\left(I-\beta F^{*}_{a\delta}\right)^{-1} E_e E_{\sigma^{e*}}\left[e^{e*}\left(a_i,x,\delta\right)\right]$$

using $\left(I-\beta F^{*}_{a\delta}\right)^{-1} E_e E_{\sigma^{e*}}\left[e^{e*}\left(a_i,x,\delta\right)\right] \triangleq e\left(a_i,x,\delta\right)$ [48], re-write (19) as

---

[48] $E_{\sigma^{e*}}\left[e^{e*}{}_i\left(a_i,x,\delta\right)\right]$ is the conditional entropy of action a, in equilibrium e when using logit errors. $E_e E_{\sigma^{e*}}\left[e^{e*}{}_i\left(a_i,x,\delta\right)\right]$ is hence the expected conditional entropy. The other terms discount and adjust for equilibrium probabilities of the action in the next period. Thus, the likelihood has the nice interpretation of being the logit probability of the sum of discounted expected payoffs in current and future periods, minus the expected conditional entropy, of any action.

$$l\left(a_i \mid x,\delta,\omega,\mu_{s_{fs}}\right)$$

$$=\frac{\exp\left(E_{\sigma_{-i}}\left[\left(\psi_i+\beta G(a)\mathrm{M}\right)\theta+\beta G(a)e\left(a_i,x,\delta\right)\mid x,\delta,\omega,\mu_{s_{fs}}\right]\right)}{\sum\limits_{a_i\in A_i}\exp\left(E_{\sigma_{-i}}\left[\left(\psi_i+\beta G(a)\mathrm{M}\right)\theta+\beta G(a)e\left(a_i,x,\delta\right)\mid x,\delta,\omega,\mu_{s_{fs}}\right]\right)} \qquad (25)$$

$$\textit{subject to} \qquad \sigma\left(a \mid x,\delta,\omega,\mu_{s_{fs}}\right)=\prod_{a_i\in a} l\left(a_i \mid x,\delta,\omega,\mu_{s_{fs}}\right),$$

$$\tilde{V}^{*}\left(a,x,\delta\right)=\mathrm{M}\theta+e\left(a,x,\delta\right)$$

where $G\left(a\right)$ is the extractor matrix that has dimensions $1\times\left\|A\right\|$ and has a 1 in the column of a, and 0s elsewhere. The (valuation operator construction) matrix $\mathrm{M}\left(x,\mu_{s_{fs}}\right)$ is found through the equilibrium mixed policy response function and the state transition function from the first stage. While AM uses the policy function in both the current and the future periods, I use the mixed policy function estimates (from the data) only in the second period to construct M. From AM, I know that the two value functions are analogues. Using(25), I write the PML estimator[49]:

$$\hat{\theta}_{PML}=\arg\max_{\vartheta\in\Theta}\prod_{d}\int_{\Omega} l\left(d \mid \hat{V}^{*},x,\delta,\mu_{s_{fs}},\omega;\vartheta\right)\mu_{\omega}\left(\omega \mid I_t;\vartheta\right) \qquad (26)$$

---

[49] The likelihood function is time homogenous with (25) defined for all observed actions. (26) is written generally to allow d to be defined for a single game, or on a panel dataset where multiple games played are played in multiple markets.

I choose to use a pseudo-likelihood-based method to maximize efficiency and ensure consistency of the standard error estimates. In an under identified model, similar to Bajari, Benkard and Levin (2007) one can instead follow Chernozukhov, Hong and Tamer (2007). Their estimator uses set identification to find parameters that describe difference equilibriums supported by the data minimizing a criterion function that penalizes violations of the best response function. The likelihood based approach is more efficient in the point-identified model, and produces precise standard errors of estimated parameters.

5.7: Estimation Algorithm and Identification

The model can be estimated in two stages. First estimate the adoption equation. The Generalized Bass Model (Bass, Jain and Krishnan, 2000) can be estimated using a range of well studied methods. Second estimate the profit functions from the restrictions implied by the theorems and lemmas presented earlier. The step-by-step estimation algorithm for the game theory based estimations is:

a. Using the estimates of the adoption model, build the state vector of the firm in every period.

b. Recover the expected (unconditional) policy function $\widehat{F}_a^*$ from the data (model and recover (18) parametrically or non-parametrically from the data). If the equilibrium selection mechanism puts point mass on an equilibrium, is the recovered expected policy function, a unique equilibrium policy function. The (i,j) entry in F_a^* is the probability that action j will be played in the next period if action vector i is played in the current period. In the FIML estimator, F_a^* is calculated for any policy function and selection probability conjectures, while the PML estimator uses the recovered policy function.

c. Recover the expected state transition function ($\widehat{F}^*_\delta$) implied by (A9), conditional on period actions from the data.

d. Build the valuation operator matrix[50], defined using

$$\widehat{M}_t = \left(I - \beta \widehat{F}^*_{a\delta}\right)^{-1} E_e E_{\sigma^{e*}} \left[\Psi_{it}\left(x_{it}, p_{it}, \delta_t, a_t \mid \mu_{s_{fs}}\right)\right]$$

$$= \left(I - \beta \widehat{F}^*_{a\delta}\right)^{-1} \widehat{F}^*_a \Psi_{it}\left(x_{it}, p_{it}, \delta_t, a_t \mid \mu_{s_{fs}}\right)$$

e. Let (1) be a matrix of 1s, of dimension $\|A\| \times 1$. Define the expected errors matrix[51]

$$\widehat{e} = \left(I - \beta \widehat{F}^*_{a\delta}\right)^{-1} E_e \left[\left(\left(F^*_a\right)_e \bullet \lg\left(F^*_a\right)_e \right)(1) \right]$$

In the unique equilibrium model $\left(\widehat{F}^*_a\right)_e = \left(\widehat{F}^*_a\right)$, while in the multiple equilibrium model $\left(F^*_a\right)_e$ has to be computed in the estimator.

f. Non-linearly maximize the PML equation (26) for a conjectured numbers of equilibriums[52]. Note the estimator involves maximizing the likelihood of the action vector, and not the individual action likelihoods[53]:

---

[50] The time subscript identifies that the matrix (due to the payoff basis) is dependent on the state of the industry, and hence changes over time.

[51] For very small true values of $\Pr\left(a, x, \delta\right)$, the first stage estimates may suggest $\Pr\left(a, x, \delta\right) = 0$. In this case $\lim_{h \to \infty} h \ln(h) = 0$ implies $E_e E_{\sigma^{e*}_{-i}} \left[e^{e*}{}_i \left(a_i, x, \delta\right)\right] = 0$.

$$l\left(a \mid x,\delta,\omega,\mu_{s_{fs}}\right)=\prod_{i\in I} l\left(a_i \mid x,\delta,\omega,\mu_{s_{fs}}\right)$$

$$=\prod_{i\in I}\frac{\exp\left(E_{\sigma_{-i}}\left[\left(\psi_i+\beta G(a)\widehat{\mathrm{M}}_i\right)\theta+\beta G(a)\widehat{e}\left(a_i,x,\delta\right)\mid x,\delta,\omega,\mu_{s_{fs}}\right]\right)}{\sum_{a_i'\in A_i}\exp\left(E_{\sigma_{-i}}\left[\left(\psi_i+\beta G(a')\widehat{\mathrm{M}}_i\right)\theta+\beta G(a')\widehat{e}\left(a_i,x,\delta\right)\mid x,\delta,\omega,\mu_{s_{fs}}\right]\right)}$$

$$\textit{subject to}\qquad \sigma\left(a \mid x,\delta,\omega,\mu_{s_{fs}}\right)=\prod_{a_i'\in a} l\left(a_i' \mid x,\delta,\omega,\mu_{s_{fs}}\right),$$

$$\widehat{\mathrm{M}}_i=\left(I-\beta\widehat{F}^*_{a\delta}\right)^{-1}E_e\left[\widehat{F}^*_{ae}\psi_i\left(x_{it},p_{it},\delta_t,a_t \mid \mu_{s_{fs}}\right)\right]$$

$$\widehat{e}=\left(I-\beta\widehat{F}^*_{a\delta}\right)^{-1}E_e\left[\left(\left(F^*_a\right)_e\bullet\lg\left(F^*_a\right)_e\right)(1)\right]$$

g. Increase the numbers of equilibriums in the optimizer and maximize, till there is no increase in the Bayesian Information Criterion (BIC).[54]

Multiple mappings from the independent variable and the shock to potential dependent

---

[52] $G(a)$ is the extractor matrix with a 1 in the column of a, and 0s elsewhere.

[53] Let $\Pr(a_i)$ be the probability of an action of i. We want $\Pr(a)=\prod_{i\in I}\Pr(a_i)$. In a unique equilibrium model, extant models maximize individual actions of agents. In the PML estimator I instead maximize the joint probability of the entire agent action vector, since the probability of the joint (mixed) actions is not the likelihood of the mixed actions. As:

$$\Pr(a_i,a_j \mid e_1)\Pr(e_1)+\Pr(a_i,a_j \mid e_2)\Pr(e_2)$$
$$=\left(\Pr(a_i \mid e_1)\Pr(a_j \mid e_1)\Pr(e_1)+\Pr(a_i \mid e_2)\Pr(a_j \mid e_2)\Pr(e_2)\right)$$
$$\neq\left(\Pr(a_i \mid e_1)\Pr(e_1)+\Pr(a_i \mid e_2)\Pr(e_2)\right)\left(\Pr(a_j \mid e_1)\Pr(e_1)+\Pr(a_j \mid e_2)\Pr(e_2)\right)$$

[54] I am optimistic that infinite equilibriums can be modeled using a continuous function for selection probability. The math for the infinite model should be similar with a mixed logit replacing the finite mixture logit likelihood.

variables exist in the model. If a unique mapping exists between these variables, the inversion of that relationship allows us to write the probability of seeing the data. With multiple mappings, I define the likelihood of an action by integrating over all mappings. The model is identified parametrically because of differences in the probability of an action vector, in each equilibrium regime[55]. To identify parameters within equilibriums, I set profits from not releasing any titles as zero. I cannot identify parameters of selection mechanism to scale, but recover equilibrium selection probabilities. If N equilibriums are played with positive probability, then the BIC for models with more than N equilibriums should be lower than the BIC for N equilibriums. A lower BIC with increasing N would imply that either a larger number of equilibriums do not exist for the model primitives, or certain equilibriums are not played with significant probability in the data.

6. Simulation and Results

In this section I describe simulations based on my MPNE frame work. My framework generalizes dynamic games to allow cooperation and competition both within a game, and in the selection mechanism between games. The simulations present cases where cooperation amongst firms in equilibrium selection increases firm profits. These simulations serve several purposes. First, they outline and motivate a case for trade-offs discussed prior. Second, similar to Dube, Hitsch and Manchanda (2005), the simulations illustrate which policy prescriptions are supportable as a MPNE. Third,

---

[55] Action vector probabilities must be different across equilibriums; else they are the same equilibrium.

they illustrate the effect of mixing across equilibriums on the equilibrium behavior of firms, the adoption behavior of consumers and the profits of firms. In future work, I plan to measure the rate of convergence of the estimation algorithm, and the extent of small sample bias due to sequential estimation. I will also estimate the model on market data described in section 3 and recover the true primitives.

Simulation 1 is a generalization of example 1 that focuses on the competitive advantage conferred by winning. Simulation 2 studies the effect of the market evolution on payoffs similar to the repeated prisoner's dilemma game. Simulation 3 focuses on the application of the frame-work to a game between studios. It requires firms to choose between maximizing current period profits and supporting the adoption of the new format.

For any action, model primitives imply a market outcome in the current and future periods. To capture this, I specify a payoff matrix (and in Simulation 3 an adoption function). I simulate the equilibrium behavior of firm for different selection rules, plotting the evolution of the industry and measuring payoffs to the firms. In each simulation, I specify the mean payoff and iid gumbel private information shocks. I use the payoffs to derive the equilibrium policy function, by analytically deriving fixed points in the equilibrium policy functions, conditional on a selection rule in the simulation. I then simulate per period choices of firms (using the derived policy functions). Last, I plot equilibrium behavior of players, and show the effect of the selection rule on firm payoffs.

6.1: Simulation 1: accruing advantages

Simulation 1 addresses the question of what advantages accrue in the game through equilibrium play. The firm with greater payoffs in an equilibrium may find a sustainable advantage in future equilibriums, leading to a first mover advantage.

The simulation extends Example 1 to a dynamic game. 1 chooses between {R,L}, and 2 chooses between {A, B}. Let the state of the game be the last action played $\delta = \{\{R,A\},\{L,A\},\{R,B\},\{L,B\}\}$. Suppose payoffs conditional on state are defined as:

Table 11: Payoffs in Simulation 1

			2	
			A	B
$\pi(a) \mid \delta = \{R,A\}$	1	R	30, 10	0, 0
		L	0, 0	5, 15
$\pi(a) \mid \delta = \{L,B\}$		R	5, 15	0, 0
		L	0, 0	30, 10

For 1, playing an equilibrium changes future payoffs to increase Player 1's payoffs from the same equilibrium. Hence Player 1 accrues advantages through an equilibrium. 2, on the other hand, prefers continuously switching equilibriums. Let {r,l} represent the probability of 1 playing R and L respectively, and {a,b} represent the probability of 2 playing A and B respectively. The two PSNE (written as {{r,l},{a,b}} are {{1,0},{1,0}} and {{0,1},{0,1}} (as highlighted in the table). Potential gains are significantly larger for a dominant firm. For instance, if 1 forces the players to choose {{1,0},{1,0}} in every period, then 1 will get a payoff of 30 in every period. Similarly if 2 forces the players to switch equilibriums in every period, then 2 gets a payoff of 15 in every period. Write payoffs at the equilibrium as $\{\pi_1, \pi_2\}$ where π_1 is the payoff to 1 and π_2 is the payoff to 2. Equilibriums payoffs are in Table 12.

Table 12: Equilibrium payoffs in Simulation 1

Equilibrium	Profit vector $\{\pi_1, \pi_2\}$
{{1,0},{1,0}} always	{30, 10}
{{0,1},{0,1}} always	{30, 10}
Continuous switching between {{1,0},{1,0}} and {{0,1},{0,1}}	{5, 15}

6.1.1: Results from Simulation 1

The equilibrium selection rule most favorable to player 2, involves switching (alternating) between the equilibriums[56]. I simulate the second strategy for equitable payoffs, where to make equal profits the studios can decide to play the last PSNE played with probability p and the other PSNE with probability 1-p. As the equilibrium coordination probability p increases, the average probability, and the variance of the probability of playing the most favored equilibrium for any player, remains constant (see Figure 7).

---

[56] For instance, firms may benefit from switching if consumers are variety seeking; they may see higher profits when changing strategies.

Total payoffs to the players vary with coordinating probabilities, with player 1 benefiting from larger probabilities, and player 2 benefiting from smaller values. Figure 8 plots total payoffs received by either player over 100 periods. With p=0.33, the two players receive identical expected payoffs of 13.33. In the data, we may not necessarily see firms play the equitable (profit sharing) selection rule. If player 1 wields more market power, then a rule preferred by player 1 maybe played in the data.

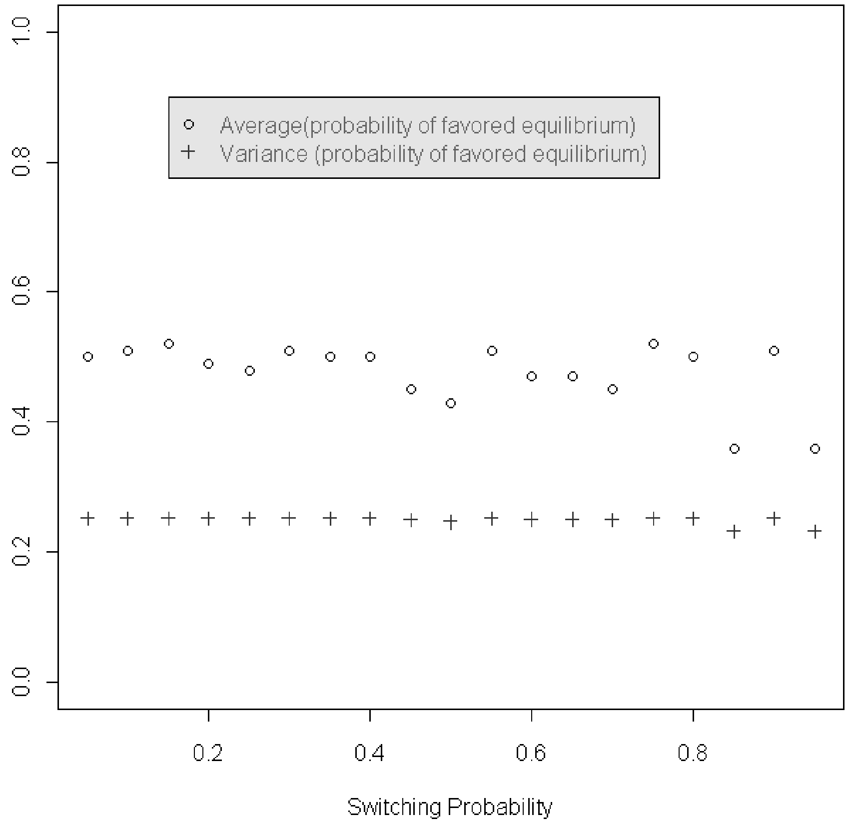

Figure 7: Probability of the favored equilibrium

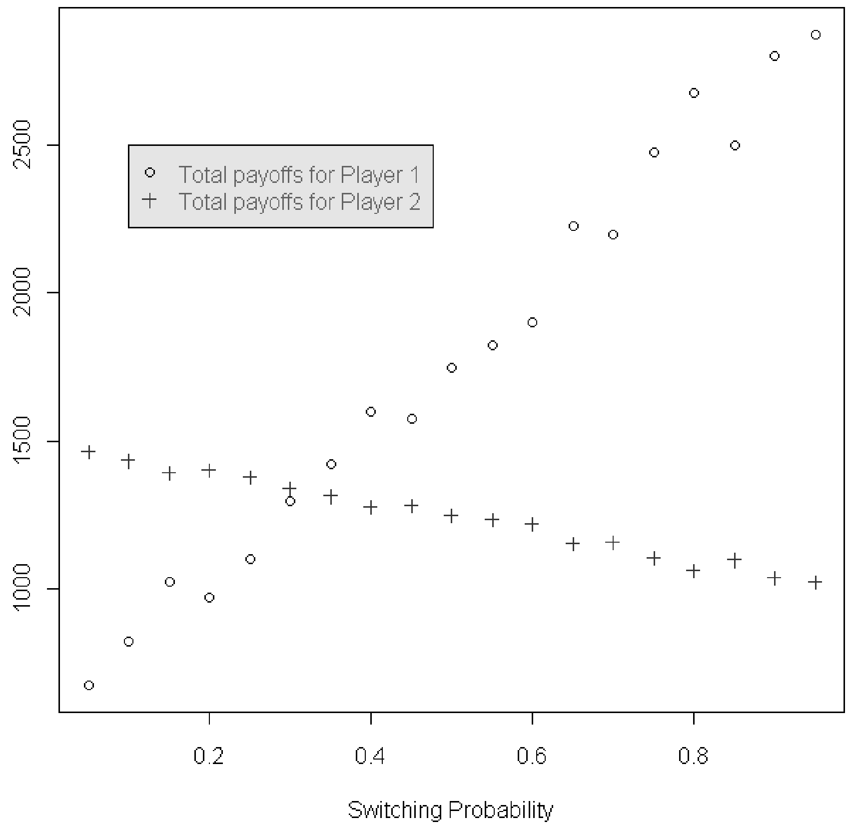

Figure 8: Total payoffs for player 1 and 2

6.2: Simulation 2

Simulation 2 tests responses of firms to different levels of adoption. Suppose there are different rules for different levels of adoption: when adoption is low, an equilibrium selection rule is used to select between equilibriums, and an alternative selection rule is used when adoption is high. Consider a game in which player 1 chooses between {R, M and L}, and player 2 chooses between {A, B and C}. Table 12 lists payoffs from the equilibrium.

Table 13: Payoffs in Simulation 2

		2		
		A	B	C
1	R	10, 10	0, 25	0, 35
	M	25, 0	50, 25	0, 10
	L	35, 0	25, 0	10, 50

The two PSNE are at {M, B} and {L, C}. The two equilibriums lead to asymmetric non pareto-dominant payoffs: 1 prefers {M, B} and 2 prefers {L, C}. Both players have asymmetric min-max punishment strategies to deter deviations. 1 can punish 2 to receiving 25 forever by playing M, keeping the equilibrium at 1's favored PSNE. 2 can punish 1 to receiving 10 forever by playing C, keeping the equilibrium at 2's favored PSNE.

6.2.1: Results from Simulation 2

Suppose two adoption levels, la and ha, with a different probability {lape1, hape1} of playing {M, B}. Lape1 characterizes the probability of playing {M, B} when the adoption level is la, and hape1 when the adoption level is ha. Let the adoption level change with uniformly increasing probability over time. In Table 14, I report profits for different lape1 and hape1 values. With increasing pe1 probability, firm 1 sees larger payoffs. If firm 1 has greater strategic power, potentially from the stronger punishment strategy in the game, then higher values of pe1 will be chosen. Conversely, if firm 2 exerts greater power, then lower values of pe1 would be chosen, and the payoffs of firm 1 reduced. In the frame-work, to model market power arising from asymmetric payoffs, I allow equilibrium selection probabilities to be a function of the total adoption till date, and the last actions and states of firms.

Table 14: Equilibrium payoffs as function of selection probabilities

lape1	hape1	Total payoff 1	Total payoff 2
0.1	0.9	1712	4555
0.3	0.7	2354	4153
0.5	0.5	3012	3742
0.7	0.3	3546	3409
0.9	0.1	4190	3006

6.3: Managing format transitions

Simulation 3 formalizes arguments made prior in the paper regarding the transition between VHS and DVD, to the mathematical frame-work. The aim of the simulation is to (a) show the frame-work, and in particular selection mechanism, is general enough to allow for discussed effects, and (b) to show the importance of the selection mechanism in defining the outcome of the game. Hence the simulation specifies mean payoffs to the studios taking an action: the specified payoff matrix is a reduced form capture of the market outcomes of studio actions. Similar to the frame-work, I assume agent payoffs also receive private information shocks, revealed prior to choosing actions in a period, and show the effect of different selection rules.

In simulation 3, I assume 2 studios choose between 3 levels of actions. For the discussion, assume these actions reflect three DVD availability levels, low (LA), medium (MA) and high (HA). These actions could instead reflect different mean prices or any joint pricing and availability decisions. Next suppose Table 15 reflects the action payoffs with $\{m_1, m_2\}$, where $profits(i,t) = m_i \bullet Adopt\,index(t)$. This corresponds to a market share based model, where payoffs would be similar with the adoption index reflecting the total consumer base interested in purchasing DVDs in a period, and the matrix payoffs, a reweighting of the market share of the studios. I assume the adoption index follows:

$$Adoption\,index\,(t) = \gamma_0 + \gamma_1 Adoption\,index\,(t-1) + \gamma_2 HA_i + \gamma_3 MA_i + \gamma_4 HA_j + \gamma_5 MA_j$$

with $\{\gamma_0 = 0.8, \gamma_1 = 0.15, \gamma_2 = 0.05, \gamma_3 = 0.15, \gamma_4 = 0.05,\}$.

Table 15: Availability and payoffs

		2		
		HA	MA	LA
1	HA	5, 4	3, 2	5, 4
	MA	2, 3	4, 5	6, 3
	LA	4, 5	3, 6	5, 5

6.3.1: Industry conduct in Simulation 3

My primary research question is motivated by the trade-off between supporting adoption for future periods, and playing the Nash Equilibrium that leads to the highest current payoffs, when studios have asymmetric payoffs from promoting adoption. Hence I specify a game where the equilibrium at {HA,HA} leads to a higher mean adoption index in future periods, and hence is preferable to both players in the long run, but where player 2 prefers {MA,MA} in the short run.

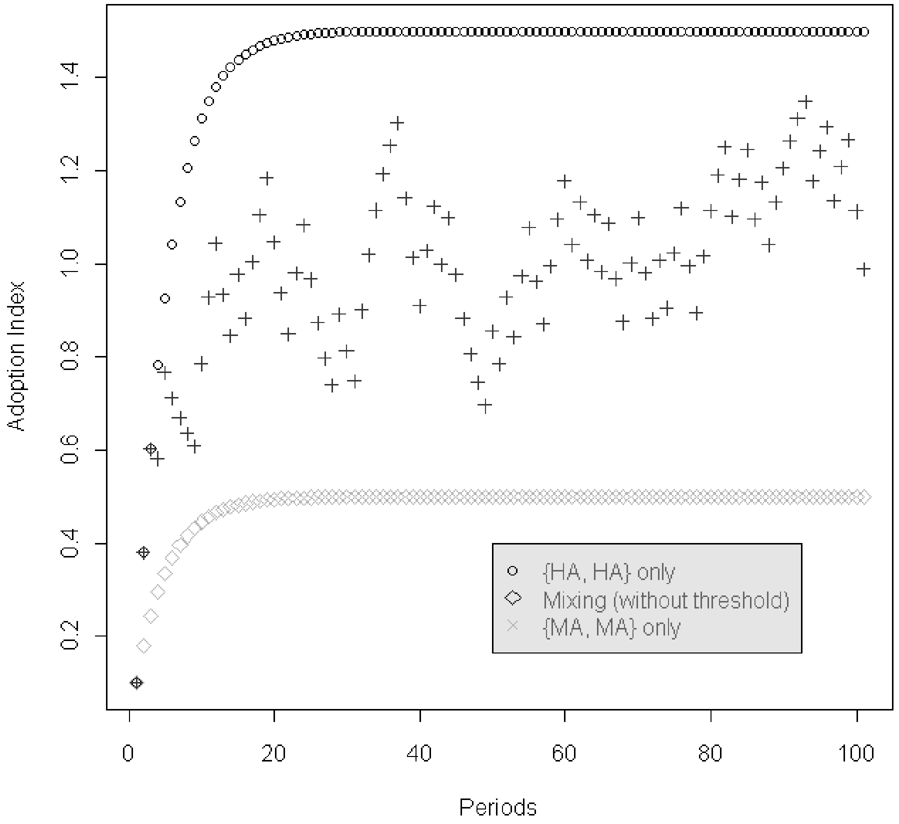

Figure 9: Adoption rates (without threshold)

Figure 9 shows the adoption levels as a function of the equilibrium chosen. Player 2 gets a lower (immediate) payoff at {HA,HA} and may entirely prefer {MA,MA} at certain discount rates. Instead players may randomly select between equilibriums to ensure greater adoption than at {MA,MA}. Further, firms may choose to change the mixing probability when the market reaches a certain adoption level (threshold), as in simulation 2. If studios initially choose equilibriums in which they price (or release DVDs) more aggressively then more consumers adopt into the model in later periods. Hence it may be profitable for firms to cooperate choosing competitive equilibriums that increase adoption while equitably sharing resulting profits. The threshold rule allows the two firms to reach higher levels of adoption without leveling off (Figure

10) at the adoption levels for {MA,MA}. Table 16 shows that both players prefer the mixing rules over playing {MA,MA}. However, Player 1 prefers the threshold mixing rule to the pure mixing rule, whereas player 2 prefers the pure mixing rule. For a high enough discount rate, player 2 prefers the pure mixing rule to {HA, HA}. Table 16 lists total payoffs from different selection mechanisms.

Table 16: Payoffs from different selection mechanisms

	Player 1	Player 2	Player 1 (Discounted @0.75/period)[57]	Player 2 (Discounted @0.75/period)
{HA,HA} only	715	429	9.375	5.65
Random mixing	380	407	5.71	5.82
Random mixing (with threshold)	609	390	6.62	4.97
{MA,MA} only	144	240	2.25	3.75

---

[57] Discount rate used in the simulation was 0.99.

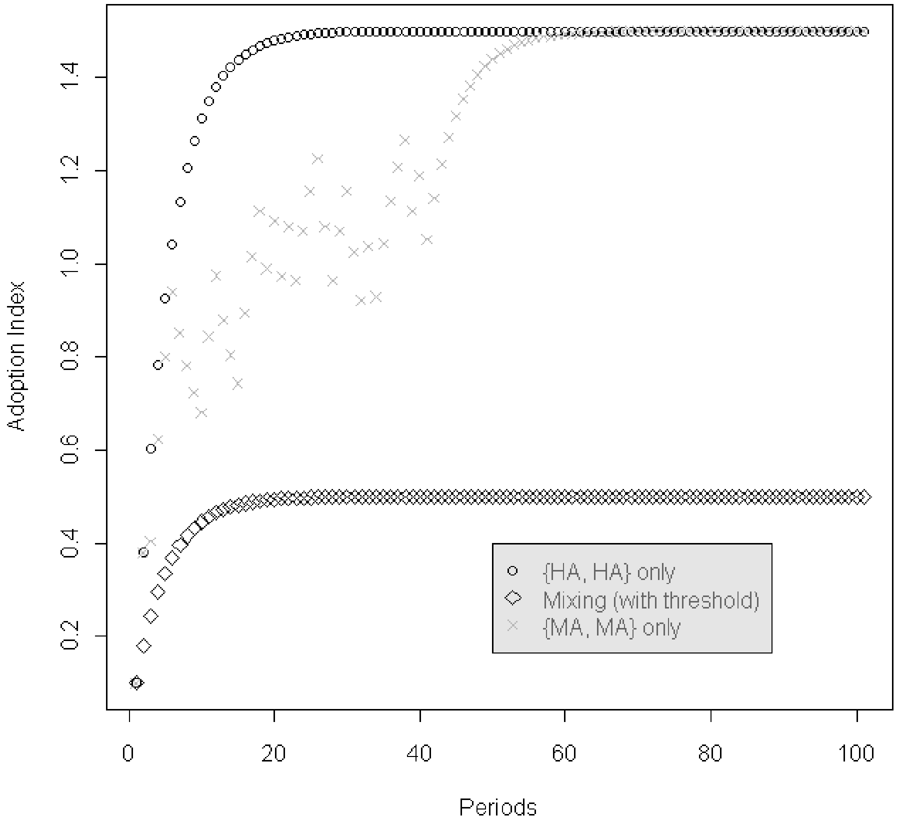

Figure 10: Adoption rates (with threshold)

The key take-away from the simulations is that studio actions (industry support for the format), and hence adoption and studio profits, differ across equilibriums. The observed actions in a dataset depend on the selection rule used to choose between equilibriums. These selection rules may be chosen based on a number of factors including the market power of different studios and difference in payoffs across equilibriums. While firms thus have incentives to ensure that they use appropriate selection rules, a priori it is difficult to specify the selection rule applicable to a particular market.

In the paper, to model equilibrium selection, I develop a frame-work that extends the extant MPNE frame-work. My specification of the selection rule allows for dependencies on past period actions and states, and admits each case simulated and discussed. The describe estimators account for equilibrium selection and remain consistent when different equilibriums are played in the data, recovering both model primitives and the selection rule, while nesting the base case of a unique equilibrium being played in the data.

7. Conclusion

Essays 1 and 2 study title specific managerial issues, and focus on forecasting demand, timing release and pricing. Essay 3 studies format and studio level managerial issues, looking at the allocation of an assortment/portfolio of titles of a studio across formats. In essay 3, I model the release and pricing of new titles on DVD and VHS at the end of the VHS technology lifecycle. In my model, studios are forward looking and strategically use release timing and pricing strategies to influence adoption. Studios face conflicting short-term and long-term profit interests when deciding pricing and assortment in the old VHS and new DVD format. Competitive dynamics are especially important, when studios have to both coordinate with rivals as well as face the temptation to free-ride off rivals. The resulting trade-offs are examined using a dynamic model that calculates the total value of a strategic choice, including its effect on future demand. Thus, I endogenize the efforts of firms to drive adoption of newer formats, and study the evolution of formats in competitive markets with substitution and complementarity effects. To model these issues in format adoption, I develop estimation methods for a dynamic Markov Perfect Nash Equilibrium (MPNE) with multiple equilibriums. As firms could coordinate strategies to drive adoption or free ride, strategies in my model could be strategic substitutes or strategic complements, leading to the multiple possible equilibriums.

I describe two estimators: an efficient but computationally expensive full information maximum likelihood estimator and an inefficient but computationally inexpensive pseudo-likelihood estimator. I limit the scope of the paper to a description of these estimators. In future work, I will compare these estimators to draw inferences on the dynamics of cooperative and competitive behavior in the industry.

There are several directions in which this work can be enhanced in future research. First, the FIML and PML estimators can be compared for convergence and efficiency, for different model primitives. The model can be extended to relax the full information assumption (no learning) implicit in the equilibrium selection mechanism. The general framework described is broad enough to be applied to other problems in which firms must make trade-offs between competitive and cooperative decisions, in a dynamic setting. Hence, substantively the study can be extended to other home video and theatrical channels, and additional adoption decision variables (eg: advertising).

BIBLIOGRAPHY

Aguirregabiria, V., and P. Mira (2002), "Swapping the Nested Fixed Point Algorithm: A Class of Estimators for Discrete Markov Decision Models," *Econometrica* 70, no. 4 (July): 1519-1543.

Aguirregabiria, V. and P. Mira (2007), "Sequential Estimation of Dynamic Discrete Games," *Econometrica*, 75 (1), 1-53

Ainslie, A., X. Dreze, and F. Zufryden (2005), "Modeling Movie Life Cycles and Market Share," *Marketing Science*, 24 (3), 508-17

Aumann, R. J. (1974) "Subjectivity and correlation in randomized strategies" *Journal of Mathematical Economics* 1, no. 1: 67-96.

Bajari, P., C. L. Benkard, and J. Levin (2007), "Estimating Dynamic Models of Imperfect Competition," *Econometrica*, 75 (5), 1331-70

Basu, A. T. Mazumdar, and S. P. Raj (2003), "Indirect network externality effects on product attributes," *Marketing Science* 22, no. 2: 209.

Beresteanu, A., I. Molchanov, and F. Molinari (2008), "Sharp Identification Regions in Games," *CEMMAP Working Paper CWP15/08*, Institute for Fiscal Studies, UCL.

Binmore, K., and L. Samuelson (1999) "Evolutionary drift and equilibrium selection," *Review of Economic Studies*: 363-393.

Chernozhukov, V., H. Hong, and E. T. Tamer (2006) "Estimation and confidence regions for parameter sets in econometric models," *working paper,* SSRN.

Chiou, L. (2007), "The Timing of Movie Releases: Evidence from the Home Video Industry," *working paper*, Occidental College, CA, USA

Doraszelski, U. and A. Pakes (2006), "A Framework for Applied Dynamic Analysis in IO," *working paper*, Harvard University, MA, USA

Dorazselski, U. and M. Satterthwaite (2007), "Computable Markov-Perfect Industry Dynamics: Existence, Purification, and Multiplicity," *working paper,* Harvard University, MA, USA

Dube, J. P., G. J. Hitsch, and P. Manchanda. 2005. An empirical model of advertising dynamics. *Quantitative Marketing and Economics* 3, no. 2: 107-144.

Dube, J.P., G. J. Hitsch, and P. Chintagunta (2008) "Tipping and Concentration in Markets with Indirect Network Effects", *working paper,* SSRN.

Einav, L. (2003), "Not All Rivals Look Alike: Estimating an Equilibrium Model of the Release Date Timing Game," *working paper*, Stanford University, CA, USA.

---- (2007), "Seasonality in the U.S. Motion Picture Industry," *RAND Journal of Economics*, 38(1), 127-45.

Elberse, A. and F. Oberholzer-Gee (2007), "Superstars and Underdogs: An Examination of the Long Tail Phenomenon in Video Sales," *working paper*, Harvard University, MA, USA.

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 (6), 638-61.

Ericson, R. and A. Pakes (1995), "Markov-Perfect Industry Dynamics: A Framework for Empirical Work," *Review of Economic Studies*, 62 (1), 53-82.

Fisher, Franklin M. (1989) "Games Economists Play: A Noncooperative View," *The RAND Journal of Economics* 20, no. 1 (Spring): 113-124.

Gupta, S., D. C. Jain, and M. S. Sawhney (1999), "Modeling the evolution of markets with indirect network externalities: An application to digital television," *Marketing Science*: 396-416.

Hauser, J., G. J. Tellis, and A. Griffin. (2006) "Research on innovation: A review and agenda for marketing science," *Marketing Science* 25, no. 6: 687.

Hennig-Thurau, T., V. Henning, H. Sattler, F. Eggers, and M. B. Houston (2007), "The Last Picture Show? Timing and Order of Movie Distribution Channels," *Journal of Marketing*, 71 (4), 63-83.

Jia, P. (2008) "What happens when Wal-Mart comes to town: An empirical analysis of the discount retailing industry," *Econometrica* 76, no. 6: 1263-1316.

Krishnan, T. V., F. M. Bass, and D. C. Jain (1999), "Optimal pricing strategy for new products" *Management Science*: 1650-1663.

Lehmann, D. R. and C. B. Weinberg (2000), "Sales through Sequential Distribution Channels: An Application to Movies and Videos," *Journal of Marketing*, 64 (3), 18-33.

Luan, J. and K Sudhir (2007), "Optimal Inter-Release Time Between Sequential Products: Application to Theatrical Movies and DVDs," *working paper*, Dartmouth College, NH, USA.

Nair, H. (2007) "Intertemporal price discrimination with forward-looking consumers: Application to the US market for console video-games," *Quantitative Marketing and Economics* 5, no. 3: 239-292.

Nair, H., P. Chintagunta, and J. P. Dube (2004). "Empirical analysis of indirect network effects in the market for personal digital assistants", *Quantitative Marketing and Economics* 2, no. 1: 23-58.

Pakes, A. and P. McGuire (2001), "Stochastic Algorithms, Symmetric Markov Perfect Equilibrium, and the 'curse' of Dimensionality," *Econometrica*, 69(5), 1261-81.

Prasad, A., B. Bronnenberg, and V. Mahajan (2004), "Product Entry Timing in Dual Distribution Channels: The Case of the Movie Industry," *Review of Marketing Science,* 2 (2004), Article 4.

Seim, K. (2006), "An Empirical Model of Firm Entry with Endogenous Product-Type Choices," *RAND Journal of Economics*, 37(3).

Swami, S., J. Eliashberg, and C. B. Weinberg (1999), "SilverScreener: A Modeling Approach to Movie Screens Management," *Marketing Science*, 18 (3), 352-72.

Tamer, E. (2003) "Incomplete Simultaneous Discrete ResponseModel withMultiple Equilibria," *Review of Economic Studies,* Vol 70, No 1, pp. 147-167.

Vitorino, M. (2008), "Empirical Entry Games with Complementarities: An Application to the Shopping Center Industry", *working paper*, University of Chicago, IL, USA

Zeileis, A. (2006), "Object-oriented Computation of Sandwich Estimators", *Journal of Statistical Software***,** 16(9), 1–16.

APPENDIX

Appendix 1: Non-Stationary MPNE Framework

This appendix describes the model primitives and resulting equilibrium in Chapter 2. It shows the conditions for identifying the non-stationary MPNE and discusses the flexibility of the model when describing multiples equilibriums played in the data.

1.1: Model Description

Following prior empirical work (Doraszelski and Pakes, 2006), we restrict our attention to symmetric and anonymous equilibriums. A set of functions, is symmetric if $f_i\left(x_i, x_{-i}, \delta_i, \delta_{-i}\right) = f_j\left(x_i, x_{-i}\delta_i, \delta_{-i}\right), \forall i, j$. Hence we abstract from the identity of the agent in the payoff function. The function, $f\left(\bullet\right)$is anonymous if $f\left(x_i, x_{-i}, \delta_i, \delta_{-i}\right) = f\left(x_i, x_{perm(-i)}, \delta_i, \delta_{perm(-i)}\right)$, where $perm(-i)$ is any permutation of the indices of other studios. Note that symmetry and anonymity restrictions do not assume that studios are identical, but instead that studio and product differences are observed. Reduced-form game payoffs for all agents at equilibrium are a function of its characteristics, state vector and competitive set.

We assume model primitives are common knowledge to potential entrants and incumbents:

$$\left\{\pi_{it}(x_{it}, p_{it}, \delta_t, a_t, :), E_t, \Psi_t(\delta' \mid \delta, a_t, :), \nu_{it}, \beta\right\}_{(i, x_i, p_i, t, \delta, a) \in I \times \aleph \times \mathrm{P} \times T \times \Delta \times A} \tag{27}$$

We describe these objects and then present the assumptions required for the model. The state space is $I \times \aleph \times \mathrm{P} \times T \times \Delta$, where $I \in \mathbb{Z}_+$ is the set of all titles ever released in movie theaters, $\aleph$ is the Cartesian product of observed movie characteristics, $\mathrm{P} \subset \mathbb{Z}_+^\infty$ is the set of prices, $T \in \mathbb{Z}$ is the set of time periods over which the game is played, $\Delta \subset T^\infty \times \mathrm{P}$ is the state vector describing announced release dates and prices and $A \subset T^\infty \times \mathrm{P}$ is the announcement vector[58]. $\left\{\Delta^C, \Delta^A\right\}$ is a partition of Δ such that Δ^C is the set of continuation states and Δ^C is the set of absorbing states. $\pi_{it}(x_{it}, p_{it}, \delta_t, a_t, :)$ is the profitability of title i, in time period t, when it has characteristics $x_{it} \in \aleph$, price $p_{it} \in \mathrm{P}$, δ_t is the state vector at time t and a_t is chosen by agents. E_t is set of all potential entrants, and $\Psi_t(\delta' \mid \delta, a, :)$ is the transition function that determines state transitions. As mentioned previously, we assume this function is time-varying. We adopt the convention of using primes to denote subsequent period variables, e.g. δ' to denote δ_{t+1}, δ'' to denote δ_{t+2}. In an abuse of notation, we also use agent subscripts to

---

[58] The lack of an announcement, or its withdrawal, is encoded as the origin.

denote the partition of the state and action vector describing an agent, e.g. δ_i to denote the state of agent i.

Incomplete information models simplify the analysis of the equilibrium (Seim, 2007), and are more likely to accurately represent the industry (e.g. given the non-standard contracts for sharing revenue and for deciding the promotional expenditure). Hence, similar to entry models, we assume that prior to making a release decision studios receive a private payoff shock ν_{it}, drawn independently over time, from a distribution $G_\nu(\bullet | \delta_t, x_{it})$ with support on $\mathbb{R}^{|E_t|}$, and ν_t the vector of private shocks for all titles in period t. Private information shocks (in a timing game) describe payoffs variations from different announcements and capture seasonal and studio specific differences in release costs across periods, including changes in the costs of advertising, promotional expenditure and the manufacturing cost for the DVD. We formalize this intuition in (T1) in section 4.3. Our formulation of private information allows shocks to be correlated across the industry, and heteroskedastic over time. Finally we specify a discount factor β. We impose further restrictions on the model primitives.

(A1) The state space is finite:

$$\left(I < \infty; x < \infty, \forall x \in \aleph; p < \overline{P} < \infty, \forall p \in \mathrm{P}; T < \infty; \delta < \infty, \forall \delta \in \Delta \right)$$

(A2) Profits are bounded ($\underline{\pi} < \sum \pi_{it}(\bullet) < \overline{\pi}$). Profits accrue post release (entry) of the title.

(A3) Studios discount future payoffs $\beta \in (0,1)$.

(A4) Private information appears additively in profit function $\pi_{it}(x_{it}, p_{it}, \delta_t, a_t, \nu_t, :) = \pi_{it}(x_{it}, p_{it}, \delta_t, a_t, :) + \nu_{it}$. F_ν is distributed absolutely continuous to the Lebesgue measure.

(A5) State transition, follows a non-stationary first order Markov process with non-homogenous transition function $\Psi_t(\delta' \mid \delta, a, :)$. In general, future states are a time varying stochastic function of past states and actions $\mu_{st}(\delta_{t+1}) = f_t(\delta_t, a_t), s.t. \forall \{\delta' \in \Delta, \delta \in \Delta^C\} \exists a$ with $\mu_{st}(\delta') = f(\delta, a) > 0$.

Assumption (A1) stipulates the finiteness of the state space. First, as we restrict our attention to titles released in movie theaters and ignore direct-to-DVD sales, our set of agents is always finite. Second, in practice, characteristics of a movie have finite range. Third, perishability implies that in any time period, a firm only considers a finite number of future periods for release. Correspondingly, we restrict the decision vector and action vector of potential entrants to be finite. Assumptions (A2), (A3) and (A4) are features of commonly-used profit functions in empirical studies. The second part of (A2) is appropriate in a game of release timing where the agent can only make profits post release of the title.

Assumption (A4) describes the stochastic monotonicity and continuity requirements on payoffs, fundamental to the existence of MPNE. Relaxing distributional assumptions on private information increases the number of mixed MPNE supported in the model. Assumption (A5) requires the state of the world, dates and prices chosen by studios in every week, to evolve in a first order Markov process. The evolution of the next period's states, conditional on the actions and states of agents in the current period, is stochastic. In keeping with extant papers, we require states in future periods to be accessible from continuation states. In our application, the state space is a history of past actions, and evolves deterministically conditional on past states and actions. (A4) and (A5) are implicitly equivalent to requiring additive separability of controls and errors, and conditional independence. Finally our state space is finite (unlike Ericsson and Pakes (1995), who study a problem with an infinite state space), and hence, we do not require agent payoffs to be bounded at the extremums of the state space.

In appendix 2, we describe an extension to the model that allows us to model price as a continuous variable. Thus, we extend the model to both continuous and discrete controls, from discrete controls. As the extension is computationally more expensive and requires additional assumptions (A11) and (A12), we maintain the assumption of discrete controls for the remainder of this paper. Assumptions (A1 – A5) lead to the following lemma:

Lemma (L1): Generically, the best response function is a unique mapping from $I \times \Delta \times T \to A$.

Proof: (L1) follows naturally under assumptions (A1), (A2), (A3), (A4), (A5) and the solution concept of a Perfect Bayesian Equilibrium (PBE). In a PBE, a studio considers expected profits from each strategic choice. The expectation on the profit function includes probabilities on the decisions of incumbents and potential entrants in future periods, described by the transition kernel specified in (A5). When taking expectations, indifference between two actions, making the best response function a correspondence occurs on a set of measure zero due to the continuity restrictions imposed in (A4).

Corollary (CL1):

If a firm is in a continuation state, $\delta_i \in \Delta_i^C$, then all future states of the firm have positive probability. $\int_{a_t} \mu_s\left(\delta_{t+1} \mid a_t, \delta_t\right) \mu_a\left(a_t\right) > 0$

Proof: (A4) implies that from each state, all actions are played with positive probability. (A5) implies that generically, all future states are achievable from continuation states.

For a particular equilibrium, (L1) implies a unique mapping from a combination of the state vector and observables to a future state. In our model, the state vector is the history of past agent actions, making the evolution of the state space conditional on agent actions, purely deterministic. In general, if agent actions lead to stochastic state changes, then agent beliefs in the PBE are rational, as (L1) implies (A5).

Our non-stationary formulation extends extant frameworks of dynamic games, to allow for multiplicity of equilibriums played in the data. If a particular equilibrium is played in the data, then the specification of the stationary MPNE (with a time homogenous transition kernel) is complete without specifying additional beliefs on equilibrium arbitration. If multiple MPNEs are possible in a model, then different equilibriums may lead to different transition kernels. The presence of multiple equilibriums in a dataset means that the transition kernel is time-varying: equilibrium choice in a particular period determines state transitions in the period. The non-homogenous transition kernel defined in (A5) is agnostic on the source of the non-stationarity and hence is broad enough to allow for the effect of multiple equilibriums on state transitions. A stationary Markov model assumes that best responses depend only on the current state of the agents. Doraszelski and Satterthwaite (2007), prove the existence of a MPNE, and under certain conditions, the existence of a Pure Strategies MPNE. Assuming a stationary MPNE, they write the choice value function as

$$V(x_t, a_t, \delta_{it}, \upsilon; \theta) = \pi(x_t, a_t, \upsilon; \theta) + \beta E_{\delta_{t+1}|\{\delta_t, a_t\}} E_{\upsilon} V(x_t, a_t, \delta_{t+1}, \upsilon; \theta) \qquad (28)$$

where expectations on future value functions are taken over possible next period states, using the transition matrix.

A stationary Markov strategy for a studio is a function $\sigma_i : \Delta \times \nu \rightarrow A$. A stationary Markov strategy profile σ is a set of stationary Markov strategies for each studio in a period. The necessary and sufficient equilibrium conditions in a stationary MPNE are

$$V\left(\delta; \sigma\right) \geq V\left(\delta; \hat{\sigma}_i, \sigma_{-i}\right), \forall i, \delta, \hat{\sigma}_i \in I, \Delta, \Sigma \qquad (29)$$

A non-homogenous first order transition matrix requires us to rewrite the choice value function. We write (29) as a period-specific choice value function, taking expectations over the next period choice value functions using the current periods' transition matrix

$$\begin{aligned} V_t(x_t, a_t, \delta_t, \upsilon; \theta) &= \pi_t(x_t, a_t, t, \upsilon; \theta) \\ &\quad + \beta E_{\delta_{t+1}|\{\delta_t, a_t\}} E_{\upsilon} V_{t+1}(x_t, a_t, \delta_{t+1}, \upsilon; \theta) \end{aligned} \qquad (30)$$

(29) specifies a time invariant choice value function, while (30) specifies a time-varying choice value function[59]. Time-varying choice value functions, particularly when lacking estimates of the transition matrix, cannot be analyzed using extant methods without arbitrary restrictions on V_t. As current decisions are affected by future seasonality, for instance to control for the effect of seasonality one would need to make V_t a function of future periods.

A non-stationary Markov strategy for a studio is a function $\sigma_{it} : \Delta \times \nu \rightarrow A$. A non-stationary Markov strategy profile σ_t is a set of non-stationary Markov strategies for each studio in period t. In a non-stationary MPNE, the necessary and sufficient equilibrium conditions are

$$V_t\left(\delta;\sigma_t\right) \geq V_t\left(\delta;\hat{\sigma}_{it},\sigma_{-it}\right), \forall i,\delta,t,\hat{\sigma}_{it} \in I,\Delta,T,\Sigma \tag{31}$$

---

[59] Blackwell's theorem does not apply to the general class of non-stationary Markov Perfect Nash Equilibriums. For instance, consider an infinite period game in which the market grows faster than the discount rate. In our problem, we assume an upper bound on the profit function and requirement the existence of an absorbing state. These two conditions, used with backwards induction arguments, guarantee the existence of such a function.

1.2: Proof of Existence of a Non-Stationary MPNE

To prove the existence of a non-stationary MPNE, we assume:

(A6) Let t_e be the time an agent enters the game. Agent payoffs $\pi\left(t, t_e, \bullet\right)$ are a strictly decreasing function of t $\pi\left(t, t_e, \bullet\right)$ with $\exists \infty > \bar{t} > t_e, s.t. \forall t > \bar{t}, \pi\left(t, \bullet\right) < 0$.

Assumption (A6)[60] is stronger than the standard waiting costs assumptions in extant frameworks due to underlying market growth; firms in markets growing fast enough may prefer to defer release indefinitely despite convex waiting costs. Estimates of inter release perishability from the market share model in our application, support (A6) with longer inter release periods leading to a sharp decline in movie appeal.

---

[60] Without (A6), the existence of the equilibrium can be proven by defining the extended model using countable states. Such a model with countable states is harder to identify than the extended model presented.

Lemma (L2): Agents have a finite planning horizon $F_M < \infty$, where M is the number of periods in which profits accrue, post release.

Proof: Define $f_t^M = \min_n, s.t. \beta^n M \overline{\pi} < \sum_{i=0}^{M-1} \pi_{t+i}$ where $\overline{\pi} = \sup_{m>t+n} \pi_{im}$. The left hand side of the inequality decreases geometrically indicating $f_t^M < \infty$. By construction, $t' > t + f_t^M, \beta^{t'-t} \sum_{i=0}^{M-1} \pi_{t'+i} < M\overline{\pi} < \sum_{i=0}^{M-1} \pi_{t+i}$ as (A6) implies a decreasing profit function. If firms never receive more profits in periods beyond f_t^M than in the current period, then they have a finite planning horizon (f_t^M) for the current period. $F_M = \min_{t \in T} f_t^M < \infty$ is the finite planning horizon for the agent across all periods.

Following Dutta and Sundaram (1994), define an extended state space, $\lambda = \{t, \delta\}$ and an extended transition matrix, $\iota(\lambda' \mid \lambda)$. Firms within a period only consider a finite number of future periods for release. A game in each period can be replaced with an equivalent finite game if we drop unconsidered states (strictly dominated states) from the extended state space creating a finite extended state space. Assumptions A1 – A5, translate in the extended state space to a stationary Markov chain. The conditions in the extended model match those in Doraszelski and Satterthwaite (2007), proving the existence of a Pure Strategy Nash Equilibrium. As the extended model notation is a one to one re-parameterization of the original model with each week and state in the original model corresponding to a state in the extended model, an equivalent equilibrium exists in the original model.

While the extended state-space notation is useful for proving the existence of non-stationary MPNE, it is not helpful to the econometrician because by construction, the cardinality of non-zero elements of the state transition matrix is always larger than the number of observations in the data. Thus, the transition matrix remains under-identified in the re-parameterization as in the original model.

We cannot compare the unidentified general non-stationary MPNE with our partial information model. Hence we assume:

(A7) The integrated value function $E_{\delta'|\delta}V_t(x_t,\delta_t,\upsilon_t;\theta) = E_{\delta'|\delta}V(x_t,\delta_t,\upsilon_t,\wp_V\left(t\right);\theta')$, and the Markov kernel, $\Psi_t(\delta' \mid \delta, a_t,:) = \Psi\left(\delta' \mid \wp_\psi\left(t\right),\delta,a_t,:\right)$, are both stochastic functions of a $\wp_\psi\left(t\right)$ and $\wp_V\left(t\right)$, finite cardinality function vectors of the effect of time t. Define the augmented characteristics vector as $x_t^A = \left\{x_t,\wp_\psi\left(t\right),\wp_V\left(t\right)\right\}$.

Assumption (A7) limits the effect of time varying payoffs to a sufficient statistic of any finite cardinality, and integrates these vectors into the vector of descriptive characteristics. While the MPNE specified in section 3 is not identified due to the non stationarity of the transition function, the augmented model (defined above) is identified. Identification requirements from the data scale with the length of the sufficient statistic used. A longer sufficient statistic remains identified in population, but increases the data requirements of the empirical implementation.

(A7) is reasonable in our model, where each agent has a finite planning horizon with a finite number of payoff periods post release. The integrated value function is well approximated in our model using the projected seasonal demand in the near future. The Markov kernel, using Perfect Bayesian restrictions, in turn is well approximated by exogenous shocks in a finite number of future periods, fulfilling the second half of the requirement. In general, non-stationary MPNE may not be well approximated by these assumptions.

Additionally, we make assumptions presented in section 5.1 (assumptions on sequences of profit functions) in Weintraub et al (2007). The demand models discussed prior (logit share, nested logit share and random coefficient logit share) are consistent with these assumptions. We get

$$\lim_{n\to\infty} E_{\delta^{(n)}}\left[V_{(n)}^{A}\left(x_t^{A},:\right)-\tilde{V}_{(n)}^{A}\left(x_t^{A},:\right)\right]=0 \qquad (16)$$

where $\tilde{V}_{(n)}^{A}\left(x_t^{A},:\right)$ is the model approximation to the identified augmented model using state distribution assumptions outlined earlier, for a market with size n. Assuming light tail conditions, as specified in section 5.4 of Weintraub et al (2007), leads to the main observation in their paper that the discounted sum of differences between actual and oblivious single period profits, converges to zero.

Non-stationary MPNE that follow (A10) may not possess a recurrent class of states, a property required for many well known (and popular) extant MPNE algorithms. For instance, approximating MPNE using the method of Pakes and McGuire (2001), requires the presence of a recurrence class for the adaptive updates to converge to true values. As the long term distribution of a non-homogenous Markov process is not well defined, adaptive learning processes may never converge to the true value, when modeling a non-stationary MPNE. However, capturing the effect of time in the finite vector may allow the use of popular two-step estimation processes, such as Bajari, Benkard and Levin (2007).

Last, (A10) generalizes the convergence result for Oblivious Equilibriums to games with multiple equilibriums played in the data, and an equilibrium arbitration process on future play. A multiplicity of equilibriums played in the data can be described by a finite family of homogenous Markov kernels, driving a non-homogenous Markov process. Our non-stationary MPNE representation is a sufficient descriptor of such a game. The game admits (A10) as an appropriate assumption, if the equilibrium arbitration mechanism (refer to Aguirregabiria and Mira, 2007 for a discussion) is driven by observed strategic and descriptor variables.

1.3: Alternative Representation of Equilibrium

For ease of exposition, we focus on the partial information model. Formally, first define the iterated (non-homogenous) unconditional Markov density as $\psi_t^{\,1}\left(\delta' \mid \delta\right) = \int_a \psi_t\left(\delta' \mid \delta, a\right)\mu_a\left(a\right)$ and $\psi_t^{\,n}\left(\delta'' \mid \delta\right) = \sum_{\delta' \in \Delta} \psi_t^{\,n-1}\left(\delta'' \mid \delta'\right) \int_a \left(\psi_{t+n-1}\left(\delta' \mid \delta, a\right)\right)\mu_a\left(a\right)$ for $n \geq 2$. Within the equilibrium for a specific state vector δ, the iterated kernel $\psi_t^n\left(\delta' \mid \delta\right)$ reflects beliefs on the state distribution n periods into the future. The iterated Markov density in the partial information model $\widetilde{\psi}_t^{\,n}\left(\delta' \mid \delta\right)$ reflects agent beliefs of the evolution of the industry, n periods into the future. Within the partial information model, integrating the iterated kernel over the beliefs on the current period's state distribution $\tilde{\varsigma}\left(\delta_{UR}^t\right)$, gives us future state distributions. This distribution $\int_{\left\{\delta_{UR}^t\right\}} \tilde{\psi}_t^n\left(\delta' \mid \delta_{UR}^t, \delta_R^t\right)\tilde{\varsigma}\left(\delta_{UR}^t\right)$ is a function of past states and past actions due to the Perfect Bayesian restrictions on $\tilde{\varsigma}\left(\delta_{UR}^t\right)$.

Lemma (L3):

$$\underset{a_{it} \in A_i}{Max}\left\{EV_t\left(a_{it}, \bullet\right)\right\}$$

$$= Max\left\{\begin{array}{l} \underset{a \in A_{it}^a}{Max}\left\{E\pi_t\left(a, \bullet\right)\right\},\ \underset{a \in A_{i(t+1)}^a}{Max}\left\{\beta E\pi_{t+1}\left(a, \bullet\right)\right\}, \\ \ldots,\ \underset{\delta \in A_{i(t+F_M)}^a}{Max}\left\{\beta^{F_M} E\pi_{t+F_M}\left(a, \bullet\right)\right\} \end{array}\right\}$$

where A_{it}^{a} are actions leading to absorbing states for agent i, and $A_{it}^{c} = A - A_{it}^{a}$, and expectations are taken over the equilibrium state distribution in the period, $\tilde{\varsigma}\left(\delta_t\right)$ and private information shock in each period.

Proof: By definition the two partitions of the action space $\left\{A_t^a, A_t^c\right\}$ are mutually exclusive and collectively exhaustive. Hence we get

$$\underset{a_{it}\in A_i}{Max}\left\{EV_t\left(a_{it},\bullet\right)\right\} = Max\left\{\underset{a_{it}\in A_t^a}{Max}\left\{EV_t\left(a_{it},\bullet\right)\right\}, \underset{a_{it}\in A_t^c}{Max}\left\{EV_t\left(a_{it},\bullet\right)\right\}\right\} \tag{32}$$

By continuing in the game, a firm obtains no profits in the current period, but gains the ability to either release in the next period, or choose a different announcement strategy. In a current period, the expected choice value of releasing is the present value of profits in the next periods. Hence, we get

$$\begin{aligned}
&\underset{a_{it}\in A_{it}^c}{Max}\left\{EV_t\left(a_{it},\bullet\right)\right\} = \\
&\quad \underset{a_{i(t+1)}\in A_{i(t+1)}}{Max}\left\{\beta EV_{t+1}\left(a_{it},\bullet\right)\right\} = \\
&\quad Max\left\{\begin{matrix} \underset{a_{i(t+1)}\in A_{i(t+1)}^a}{Max}\left\{\beta EV_{t+1}\left(a_{it},\bullet\right)\right\}, \\ \underset{a_{i(t+1)}\in A_{i(t+1)}^c}{Max}\left\{\beta EV_{t+1}\left(a_{it},\bullet\right)\right\} \end{matrix}\right\}
\end{aligned} \tag{33}$$

Implicitly the iterated Markov density allows us to take expectations over candidate states in a period. At the end of the planning horizon, a firm chooses between release in that period and continuing in that game. Substituting iteratively (33) into (32) recursively until the planning horizon we get the expression of the lemma with an additional term in the choice set of the continuation value past the horizon. For each absorbing state, the continuation value is the profits from releasing the movie. Substitute the profit function for the continuation value. From (L2) we know that profits past the horizon are lower than current period release profits, and hence can never be the argmax. Hence, (32) leads to (L3).

(L3) formalizes the earlier discussion. In expectation, the search for the maximizing release announcement strategy in a period is equivalent to a search for optimal stopping points across periods. Agents seeking to maximize expected revenues post release, maximize expected payoffs from strategic choices in a period.

In our game, this equivalence as stated is not useful as a private information shock is defined for each potential action. The search for the maximizing strategy in a period is the search for the choice value of each action including the private information shock.

Theorem (T1):

Let $\widehat{\upsilon}$ be the vector of payoff shocks, of cardinality $\sum_{i=0}^{F_M - 1} \left| A_{it}^{a} \right|$, to the expected payoffs from choosing an absorbing state. These shocks are distributed with a density μ_{π}, absolutely continuous with respect to the Lebesgue measure. Then:

(i) Private information shocks υ can be considered to be a linear combination of $\widehat{\upsilon}$

(ii) A search for the optimal absorbing state across periods is equivalent to a search for the optimal strategy in a period

Proof:

The choice of an action influences the transition of states in the period. The state density n periods in the future conditional on the action vector is $\tilde{\varsigma}\left(\delta_{t+n} \mid a\right) = \sum_{\delta'} \tilde{\psi}_{t+1}^{n-1}\left(\delta'' \mid \delta'\right) \int_{\delta_t} \tilde{\psi}_t\left(\delta' \mid \delta_t, a\right) \tilde{\varsigma}\left(\delta_t\right)$. Note that payoffs in the model accrue to the firm post release. Hence, the choice value of an action $V_t(a) = \sum_{j=0}^{F_M - 1} \left(\int_{\delta \in \Delta_{i(t+j)}^{abs}} \pi_t\left(\delta_{t+j}, \bullet\right) \tilde{\varsigma}\left(\delta_{t+j} \mid a\right) \right)$ is the expectation over resulting future absorbing states, distributed $\tilde{\varsigma}\left(\delta_{t+n} \mid a\right)$. The iterated Markov kernel has full rank and can be inverted. We can write $V_t(a) = EV_t(a) + \varphi_{\upsilon t} \widehat{\nu}$, where the mean choice value of the absorbing state is perturbed by payoff shocks $\widehat{\upsilon}$. $\varphi_{\upsilon t}$ is the matrix defined through the inverse iterated Markov kernel integrated over the state density in future periods. The mapping is unique allowing us to compare with (A4) and get $\nu_t = \varphi_{\upsilon t} \widehat{\nu}$.

In (L3), consider the case where $\hat{\upsilon}$ is added to the expected choice value of each absorbing state. Using (T1i), the resulting change in expected choice value can be captured by the private shocks to the choice value of actions in the current period. Hence, we get

$$\begin{aligned}
&\underset{a_{it}\in A_i}{Max}\left\{EV_t\left(a_{it},\bullet\right)+\upsilon\left(a_{it}\right)\right\} \\
&\qquad = Max\left\{\begin{array}{l} \underset{a\in A_{it}^{a}}{Max}\left\{E\pi_t\left(a,\bullet\right)+\hat{\nu}\left(a\right)\right\},\ \underset{a\in A_{i(t+1)}^{a}}{Max}\left\{\beta E\pi_{t+1}\left(a,\bullet\right)+\hat{\nu}\left(a\right)\right\}, \\ \ldots,\ \underset{a\in A_{i(t+F_M)}^{a}}{Max}\left\{\beta^{F_M}E\pi_{t+F_M}\left(a,\bullet\right)++\hat{\nu}\left(a\right)\right\} \end{array}\right\} \qquad (34)
\end{aligned}$$

Define $\sigma_{it}^{n}:\Delta^{n}\times\nu^{n}\rightarrow A^{n}$ as the n-period non-stationary MPNE strategy of an agent, for any finite n. A n-period non-stationary Markov strategy profile σ_t^n is a set of non-stationary n-period Markov strategies in period t. The non stationary MPNE equilibrium conditions (31) can be rewritten as

$$V_t\left(\delta;\sigma_t^n\right)\geq V_t\left(\delta;\hat{\sigma}_{it}^{n},\sigma_{-it}^{n}\right),\forall i,\delta,t,\hat{\sigma}_{it}^{n}\in I,\Delta,T,\Sigma^{n} \qquad (35)$$

(35) states that the equilibrium condition of agents choosing a maximizing strategy in each period subject to the strategies of others, is equivalent to agents seeking the maximizing n period strategy subject to equilibrium n-period strategies of others.

From (L2), we know a finite planning horizon exists for the firm. Set F_M as n and substitute in(35). The equilibrium F_M -period strategies can be considered equivalently to be those leading to the payoff maximizing absorbing state in the planning horizon. Hence, the per period equilibrium conditions of the MPNE are equivalent to agents choosing the F_M period strategy which leads to the payoff maximizing absorbing state in the finite planning horizon. Intuitively, agents searching for a strategy to maximize the choice value are searching for the F_M period strategy that leads to the maximizing absorbing state.

1.4: Estimation

Similar to Bajari, Benkard and Levin (2007), to reduce computational load we assume:

(A8) The profit function, conditional on the demand function defined, is linear in unknown parameters $\pi_{it}\left(x_{it}, p_{it}, \delta_t, a_t; \theta \mid q\left(x_{it}, p_{it}\right)\right) = \Psi_{it}\left(x_{it}, p_{it}, \delta_t, a_t \mid q\left(x_{it}, p_{it}\right)\right) \bullet \theta$ where $\Psi_{it}(\bullet)$ is a finite dimension vector of “basis functions” (including polynomial and interaction terms).

(A8) allows us to approximate the payoff function locally. A violation of (A7) does not prevent estimation or affect identification of the model and the described estimation methodology is robust to the use of a non-linear specification. As observed in Bajari, Benkard and Levin (2007), having a payoff function that is linear in unknown parameters implies that the constructed value functions are also linear in unknown parameters, simplifying estimation.

To calculate the payoffs post release, we utilize a sufficient statistic for the impact of the evolution of the industry and require:

(A9) Competition in the industry is described by an industry summary statistic set (s_{fs}). There exists a consistent estimator $\hat{\mu}\left(s_{fs}\right): \hat{\mu}\left(s_{fs}\right) \xrightarrow{p} \mu\left(s_{fs}\right)$, where $\mu\left(s_{fs}\right)$ is the true distribution of the summary statistic in a future period.

Assumption (A9) is similar to assumptions made in Bajari, Benkard and Levin (2007).[61] Instead of assuming a finite parameter vector in the first stage of estimation, we assume the forecasted variables resulting from the first-stage converge to the rational beliefs of agents.[62] (A8) can be used in other two-step dynamic models to allow the first stage regression to be non-parametric, as the summary statistic set is not

---

[61] In non-stationary MPNE, (A8) implicitly requires (A7). The sufficient statistic in (A8) can only be predicted if the non-homogenous Markov process can be modeled as a homogenous Markov kernel and exogenous time varying variables.

[62] Additional rate of convergence and local smoothness assumptions are required if using a criterion function for estimation as in Bajari, Benkard and Levin (2007).

limited in scope (and may be uncountable). For instance, the summary statistic vector may include the transition kernel and policy functions defined in Bajari, Benkard and Levin (2007). (A8) also naturally follows when a consistent parametric first stage estimator is used to estimate both, the transition kernel and the policy functions as in most dynamic game estimation methodologies.

We assume the following regularity conditions:

(R1) $\theta_{ss} \in \Theta_{SS}$ is a compact subset of $\Re^{|\theta_{ss}|}$ and true value $\theta_{ss}^{0} \in \operatorname{int} \Theta_{SS}$.

(R2) The quasi-likelihood function $\Upsilon\left(\bullet \mid \mu\left(s_{fs}\right)\right)$ is uniquely maximized at θ_{ss}^{0}, and $\Upsilon\left(\bullet \mid \mu\left(s_{fs}\right)\right)$ is twice continuously differentiable in $\theta_{ss} \in \Theta_{SS}$ with probability 1.

(R1) and (R2) are common regularity conditions for quasi-likelihood estimation, met by the iid Gumbel specification of absorbing state payoff shocks in our application. Under (A8), the argmax of the quasi-likelihood function is a consistent estimator of the second stage structural parameters. From (A4), the second stage quasi likelihood function $\Upsilon\left(\bullet \mid \mu\left(s_{fs}\right)\right)$ is continuous, leading to $\Upsilon\left(\bullet \mid \widehat{\mu}\left(s_{fs}\right)\right) \xrightarrow{p} \Upsilon\left(\bullet \mid \mu\left(s_{fs}\right)\right)$. Maximizing the quasi-likelihood yields second stage structural parameters (θ_{ss}) whose variance is the sandwich estimator $A\left(\theta_{ss}\right)^{-1} B\left(\theta_{ss}\right) A\left(\theta_{ss}\right)^{-1}$, where $A\left(\theta_{ss}\right)$ is the Hessian of the log quasi-likelihood and $B\left(\theta_{ss}\right)$ is the variance of the quasi-score.

For completeness we discuss the endogeneity of observables and a method for correcting for the endogeneity bias. In our research question, endogeneity is not a concern as observables in our payoff function are lagged variables, not affected by current private information shocks. Formally, we assume:

(A10) $E_{\nu}(\nu_{it} x_{it}) = 0$

A violation of (A10) would bias the coefficients estimated due to endogeneity. We can correct for endogeneity bias in our model using two-step estimation. In the first stage, bias correction follows methods for instruments in discrete choice models. In the second stage, violation of (A9) implies that both individual agent errors and forecast errors are correlated with agent observables. Grouping the error terms leaves a single error term correlated with observables. To estimate the model, define a set of moment conditions by matching the best response with calculated best responses and interact the conditions with instrumental variables. While this method corrects for any potential endogeneity bias it is econometrically less efficient in the second stage than maximizing the best response quasi-likelihood.

Appendix 2: Modeling Title Price Choices as a Continuous Variable

In our model we make the simplifying assumption that prices are discrete. In this section, we discuss how to model a continuous strategic variable in conjunction with release date timing. Profitability in a week is given by (10) and (11). To model prices as being continuous we make two assumptions:

(A11) $q\left(x_{smwt}, p_{smw}\right)$ and $f_{smwt}\left(q\left(x_{smwt}, p_{smw}\right)\right)$ are continuous and differentiable function of p_{smw}.

(A12) $q\left(x_{smwt}, p_{smw}\right)$ and $f_{smwt}\left(q\left(x_{smwt}, p_{smw}\right)\right)$ are quasi-concave in prices.

(A11) and (A12) are common assumptions on the profit function of a firm, which allow the researcher to formulate first order optimality conditions, and guarantee the existence of a unique maximum. Hence, they are more restrictive than (A1-A5), which lead to (L1). The assumed parametric demand function in our model satisfies both (A11) and (A12).

Under (A11) and (A12), the objective function of the firm is continuous in mixed strategies. As the strategy space is bounded, mixed strategies are a compact subset of a Euclidean space. Hence by Glicksberg's Theorem (Glicksberg, 1952), a Nash Equilibrium exists in mixed strategies.

Estimators for the model can be formulated by either maximizing the probability of joint decisions or by minimizing a criterion function. The probability of seeing a joint decision can be found by using a closed form analytical solution, or through numerical simulation. To specify the criterion function, first take derivatives of (9):

$$E\frac{\partial \pi_{smwt}}{\partial p_{smw}} = \lambda q\left(x_{smwt}, p_{smw}\right) + \left[\lambda p_{smw} - \gamma x_{smwt}\right]\frac{\partial q\left(x_{smwt}, p_{smw}\right)}{\partial p_{smw}} \tag{17}$$

To obtain first order conditions, take derivatives of (10) and substitute results of (17):

$$E\frac{\partial}{\partial p_{smw}}\pi_{smw} = \sum_{j=w}^{w+M}\beta^{j-w}E\frac{\partial}{\partial p_{smw}}\pi_{smwj} + \beta^{M}E\frac{\partial}{\partial p_{smw}}\kappa_{sm}\left(w\right) \tag{18}$$

A studio maximizes profits by setting release timing dates and prices. If we assume the existence of an interior solution, conditional on a set of dates, the first order conditions allow us to specify a Lagrangian using (18) to find the maximizing price. The difference in criterion function estimators when using only discrete controls versus joint controls, is that while the best response function for the discrete levels problem requires the researcher to enumerate value functions of all strategies, for a continuous control, the researcher first enumerates possible choices of the timing variable, and then conditional on each choice of the timing variable, solve the first order conditions of the problem to find the maximizing price. Neither approach implies a decision hierarchy; the continuous controls algorithm is identically to jointly considering all joint strategic decisions. To allow for boundary solutions, one has to consider solving the relevant Kuhn-Tucker conditions instead of specifying a Lagrangian.

It is computationally expensive to simulate the probability of a release strategy and to solve using (18) in our model. The derivative of the demand function is non linear in prices and hence requires numerical minimization. As the maximizing price is found in the inner loop for every conjecture of parameters and for every choice of release dates, the computation costs outweigh benefits of implementation in our model.