Vertical Restraints in the Movie Exhibition Industry

Krzysztof Wozniak*
Federal Reserve Board of Governors

January 1, 2014

Abstract

This paper analyzes vertical restraints imposed by distributors on movie theaters in the movie exhibition industry. A structural model of industry demand and supply is estimated using a uniquely detailed panel data set of attendance and movie rental contracts collected directly from a sample of U.S. movie theaters. Welfare analysis indicates lifting the restraints would make consumers, movie theaters and distributors better off. However, distributors face a prisoner's dilemma and lack unilateral incentives to lift the restraints. The results show how vertical restraints can persist even when all market participants would be better off if they were removed.

JEL Codes: L42, L82, D83

*Payment System Studies, Division of Reserve Bank Operations and Payment Systems, Board of Governors of the Federal Reserve System, Washington, DC 20551 (e-mail: krzysztof.p.wozniak@frb.gov)

1 Introduction

Retailers are intermediaries who buy the finished product from the manufacturer and resell it to consumers. Competition between retailers is often limited by vertical restraints imposed by manufacturers, such as retail price maintenance, exclusive dealing, or exclusive territories. This paper examines the impact vertical restraints imposed on retailers have on competition and welfare. It also looks at ways in which these restraints may persist in retail markets even when manufacturers would be better off without them.

Over the past few decades, the central dispute concerning vertical restraints has been whether they have a pro-efficiency or anti-competitive impact. Most of the literature which arose assumes that manufacturers will impose restraints on retailers only if this makes them better off. Besanko and Perry (1993) show, however, that in retail markets vertical restraints can arise as a result of a prisoner's dilemma, making the manufacturers and possibly everyone else worse off. Given the ubiquity of retail markets, it is important to understand circumstances under which this can happen, as well as the implications this may have on market structure, welfare and persistence of vertical restraints.

This paper contributes to the literature by studying the U.S. movie exhibition industry (2006-2011) in which movie distributors (manufacturers) continue imposing restraints on exhibitors (retailers) even though any pro-efficiency effects have been eliminated through technology change. Finding out why distributors keep imposing these restraints has important policy implications, as vertical restraints used in the movie exhibition industry have attracted antitrust interest internationally (OECD, 1996). It also informs the greater debate over use of vertical restraints in retail industries.[1]

A separate contribution of this paper is to examine the impact of a larger *consumer choice set* on consumer welfare. Standard economic theory predicts

[1] Recent antitrust cases include *EU v. Coca Cola* [COMP/A.39.116/B2 (2005)], *Leegin Creative Leather Products, Inc. v. PSKS, Inc.* [551 U.S. 877 (2007)] and *United States v. Apple, Inc. et al.* [12-cv-2826-DLC (2013)]

that, *ceteris paribus*, giving consumers more choice has a non-negative impact on their welfare. However, retailers often cannot offer a large choice set because of high costs of stocking/displaying goods or because of capacity constraints. The movie exhibition industry is a good setting in which to study this trade-off because heterogeneous consumer tastes and uncertainty regarding product quality make the composition of the consumer choice set a crucial driver of welfare, while the number of screens in a movie theater limits how many times each movie can be shown in a given week. The type of restraints studied can be easily generalized to industries where retailers face capacity constraints, such as general retail, radio/TV scheduling, or advertising.

To answer questions posed in this paper, a structural model of industry demand and supply is constructed and estimated using a unique, detailed panel data set of moviegoer attendance and contractual arrangements collected directly from a sample of U.S. exhibitors. Consumer demand is modeled using a flexible random-coefficient logit framework which takes advantage of the rich panel structure of the data set and explicitly models consumer selection over time. By examining exhibitor screening and distributor rental pricing decisions the supply-side model explicitly estimates movie quality uncertainty. Because some of the restraints are not directly observed in the data the counterfactual simulations are designed to establish lower and upper bounds of the range of possible welfare changes given plausible restraint regimes.

The main finding of this paper is that distributors would be better off if the restraints were lifted, but they lack individual incentives to do so. When not limited by contractual restraints exhibitors can screen a wider variety of movies and adjust their schedules more quickly as they learn the true quality of movies; this results in higher attendance as well as higher profits for exhibitors and distributors. Consumer welfare gains are driven not only by higher attendance but also by the fact that with a wider variety of movies to choose from moviegoers find movies that better fit their heterogeneous tastes.

The reason distributors lack individual incentives to lift the restraints is that imposing exclusive dealing restraints for their own movies inflicts a negative externality on other distributors. By requiring that exhibitors screen its movies for extended periods distributors shrink the number of screen-time available to other distributors' movies. Conversely, unilaterally lifting restraints results in a distributor's movies being squeezed out by movies only available under exclusive dealing restraints. This situation is a real-life prisoner's dilemma and suggests a policy banning vertical restraints would increase welfare. The same welfare gain could be achieved if distributors adopted a clause in their contracts which stipulated they would not adopt restraints as long as none of the other distributors did so. These so-called Contracts that Reference Rivals (CRR) are viewed as potentially anti-competitive by the U.S. Department of Justice (Scott-Morton, 2012) - results from this paper show that they can have pro-competitive effects as well.

1.1 Contributions and Related Literature

This paper contributes to the growing body of empirical literature which takes a fully structural approach to studying the impact of vertical restraints on market equilibria: beer distribution (Asker, 2005), car dealerships (Brenkers and Verboven, 2006), movie rentals (Ho, Ho and Mortimer, 2010) or video games (Lee, 2012*a*). In addition, the richness of the data used in this paper provides a good insight into an industry where revenue-sharing contracts are used. Previous literature has examined the use of such contracts in the movie exhibition (Hanssen, 2002; Filson, Switzer and Besocke, 2005; Gil and Lafontaine, 2012) and the video rental industries (Mortimer, 2008)

This paper takes advantage of a detailed moviegoer attendance data set to model demand for movies and screening times at the theater level. While previous papers have used reduced-form models estimated using aggregate box-office revenue data (Davis, 2006*a*; Einav, 2007; Moul, 2007; Sunada, 2012), to the best of the author's knowledge this is the first paper to use a fully structural approach to capture demand in this industry. The model builds on techniques introduced in Berry, Levinsohn and Pakes (1995, henceforth

BLP) and further developed by Nevo (2001). It also uses micro-moments to aid estimation as in Petrin (2002). Following Ho (2009), prices charged by upstream manufacturers to downstream retailers are determined by a take-it-or-leave-it bargaining model.

The model addresses three challenges. First, heterogeneity in consumer preferences results in selection which needs to be taken into account, as noted by Dubin and McFadden (1984). Consumers who chose to see a movie in its first week of release are likely to have higher-than-average valuation for this movie, so conversely those who had chosen not to see it will have a lower-than-average valuation. The richness of the dataset allows the model to track a simulated sample of consumers over time, explicitly accounting for this selection process.

Second, while the demand model can be used to estimate the true quality of the movies, it cannot be used to estimate *ex ante* quality signals. Instead, to estimate these values the paper employs a fully parametrized maximum likelihood approach, using exhibitor scheduling and distributor pricing decisions for identification. This procedure controls for another source of selection: movies chosen by exhibitors for screening (and thus observed in the data set) are likely to have higher-than-average quality compared with those not screened.

Finally, the counterfactual simulation requires the knowledge of movie quality values for all movies available to exhibitors over a time period; however, the data set used for estimation includes only a subset of all movies released by U.S. distributors. This creates an additional challenge compared to papers such as Crawford and Yurukoglu (2012) and Grennan (2013), where the whole population of products is observed in the data set. Instead, quality values for movies not in the sample are simulated based on estimates from the demand- and supply-side models, taking exhibitor movie selection into account.

2 The Movie Exhibition Industry

The industry value chain consists of four stages: *production, distribution, exhibition and consumption*. The process of *production*, not analyzed in this paper, encompasses everything from the beginning until the movie is ready to be shown to paying consumers. A *distributor* owns rights to the finished movie and decides on a release strategy for multiple platforms, the first of which is showing it in movie theaters.[2] For the theatrical run the most important choice is when to release the movie, how many theaters to release it at and how much to charge theaters for screening the movie. Distributors can be split into two categories: *Majors*, the biggest studios which have their own production studios and offer a wide variety of movies, and *non-Majors*, which are smaller and offer many independently-produced movies with a narrower audience appeal.[3] *Exhibitors* are movie theater owners, controlling anything from a single theater to a nationwide chain of multiplexes. Vertical integration between distributors and exhibitors is prohibited under the 1948 *United States v. Paramount Pictures* decree.[4]

2.1 Movie rental contracts

In the United States movie rental contracts between distributors and exhibitors employ a linear pricing schedule: there are no fixed fees, and each dollar of revenue from movie ticket sales is divided between the two parties on the basis of a *revenue split* that is contracted on in advance. For example, a revenue split of 60% means the distributor gets 60% of the revenue while the exhibitor

[2]Some movies skip the theatrical run; what follows is on-demand and online services (iTunes), DVD/BluRay discs, cable and network television

[3]This paper uses the common definition of Majors as the "Big Six" distributors who are part of media conglomerates: Paramount, Warner Bros, Columbia, Walt Disney/Touchstone, Universal and 20th Century Fox

[4]These rules have been relaxed since then allowing distributors to take non-controlling stakes in exhibitors; the exception is Sony Entertainment's ownership of Loews Theaters between 1989 and 2001

gets 40%. Historically, contracts in the U.S. employed a *sliding scale*, wherein the distributor's revenue share started off high in the first week of a movie's release and fell in subsequent weeks (Einav, 2007; Gil, 2009; Gil and Lafontaine, 2012). In recent years, however, the industry has moved toward a model with a revenue split that is constant over time.[5]

An exhibitor signs a separate contract for each movie shown, and such a contract will cover the movie's entire run at a given movie theater. The contract includes the revenue split agreed upon for the movie, as well as restrictions on how flexible the exhibitor can be when scheduling the movie. Revenue splits differ between movies, with exhibitors "paying" more for blockbusters and less for niche and independent movies. The terms also differ based on how much time has passed since the movie's nationwide release: exhibitors face higher revenue splits if they want to release a movie *on the break* (the week of the nationwide release) than if they release it *on the second run* (usually four weeks or more after the nationwide release). The revenue split is the result of bargaining between the distributor and the exhibitor, and thus can differ for the same movie across different movie theaters. Renegotiations are rare.

2.2 Vertical restraints

Movie distributors employ two types of vertical restraints to influence the way exhibitors screen their movies: *no screen-sharing* and *minimum exhibition period*.

No screen-sharing stipulates that the contracted-upon movie has the exclusive use of a screen for the duration of the contract. Since this type of restraint applies to all movies, in practice this means that a movie theater can only screen as many movies as it has screens over the course of a week and

[5] Only 14% of movies in the sample were rented on sliding scale contracts - in implementation these contracts were "flattened" to allow side-by-side comparison with the rest of the sample (see Appendix 7.1.2)

cannot, for example, put on a late-night show of a horror movie on a screen that during the day shows a children's movie.

The *minimum exhibition period* is the smallest amount of time that the exhibitor is required to screen the movie for. This restraint takes the form of an unwritten agreement between the distributor and the exhibitor. Although such a contractual form is not legally enforceable, repeated interactions ensure both parties provide have sufficient incentives to respect the agreement. The value of the restraint varies between movies, and conversations with exhibitors suggest the usual duration of minimum exhibition period for movies released on the break is between two and three weeks. While the minimum exhibition period restraint will not be binding for more popular movies, exhibitors say they often have to keep less-popular movies on for longer than they would like.

The best way to view these restraints jointly is as a form of *exclusive dealing*, wherein a retailer can only carry products of one manufacturer.[6] Viewing each screen in the movie theater as a separate retailer the no screen-sharing restraint enforces exclusive dealing for one week (the shortest period of time that distributors contract over). The minimum exhibition period restraint extends the duration of the exclusive dealing period to a few weeks.

Although these restraints are standard practice across the US, UK and Australia (OECD, 1996) their "behind the scenes" nature means they rarely make the headlines, while reliable statistics is hard to come by.[7] Evidence of the no screen-sharing restraint can be gleaned from the fact that the total number of screen on which the top 100 movies play each week is, on average, equal to the total number of screens in the US. If screen-sharing was common the former number would be considerably higher than the latter. Evidence of

[6]Examples of exclusive dealing include car dealerships which only carry one brand, gas stations, Coca-Cola's agreements with fast food restaurants and movie theaters etc.

[7]It is worth noting these restraints are not used in other markets: Gil and Lafontaine (2012) describes how Spanish market exhibitors are free to choose for how long they screen a movie, while Eliashberg et al. (2009) describe a theater in Denmark which is not limited by either the no screen-sharing or minimum exhibition period restraints

the fact that most movies in mainstream release have a minimum exhibition period of at least two weeks is provided by the fact that of the top 100 highest grossing movies in 2013 not a single one recorded a drop in the number of screens in week two of more than 0.5%.[8,9]

These restraints restrict exhibitors' decisions in two ways. First, because exhibitors have a limited number of screens at their disposal they can only screen a small subset of all movies released over the course of a year.[10] Second, if an exhibitor finds that a movie's true quality is lower than the *ex ante* quality signal they may want to replace it with a higher-quality movie, but would be unable to do so until the minimum exhibition period passes. From an exhibitor's point of view the restraints are thus attendance-reducing.

2.3 Projection technology and potential pro-efficiency impact of restraints

The movie exhibition industry in the U.S. is in the process of completing a transition from analog to digital projection technology — the aim is for distributors to stop releasing analog movies by the end of 2013. Under the old, analog projection technology an additional movie reel has to be produced and shipped to the movie theater for each screen showing the movie, the cost of which is paid in full by the distributor.[11] Industry estimates put the cost of such a movie reel at $1,500 (Alimurung, 2012).

Conversations with industry practitioners suggest the restraints imposed

[8] Although many movies have three week minima their effect is impossible to see clearly on a national level because some exhibitors screening a given movie may have a three week minimum while others will have a two week minimum

[9] Data for 2013; sources: The Numbers (box office revenues), National Association of Theater Owners (number of screens in the US)

[10] This impact is similar to that achieved by an incentive used by Coca Cola Corporation, wherein it offered free coolers to small retail stores as long as they filled it exclusively with its products. The EU found this practice to be anticompetitive (Gasparon and Visnar, 2005).

[11] The exact reason for this is unclear, though conversations with industry experts suggest that this was originally designed to incentivize exhibitors to take on new, untested movies

by distributors were originally designed to better align exhibitors' incentives with their own. By imposing the restraints distributors were guaranteed a certain number of screenings for each copy of the movie they produced, preventing exhibitors from rotating their movies too quickly. Under the old projection technology the restraints may have thus had a pro-efficiency justification in that they allowed the industry to capture most of the welfare available from potential moviegoers at substantially lower movie reel production costs.[12] Digital projection technology removes such considerations, as the cost of a digital movie copy is effectively zero. It thus reasonable to say that this pro-efficiency justification of the restraints imposed by the distributor no longer applies.

Other possible justifications for the continued use of the restraints should be considered. In other markets exclusive dealing restraints can be used by manufacturers to incentivize retailers to engage more in activities which increase demand and have a positive externality, for example customer service or advertising. In the movie exhibition industry, however, customer service has very little impact on demand, while advertising is fully paid for by distributors. In addition, since the restraints only apply to a contracts negotiated when the movie is ready to be screened, they do not reduce uncertainty during the production process.

3 Data

This paper uses two primary types of data: attendance data, which reflects the realized demand in the market, and contract data, which contains revenue split information on movie rental and restraints imposed by distributors on exhibitors. These are supplemented with additional data sources.

Attendance Data Attendance figures were collected directly from five movie theaters, each of which can be thought of as having a local monopoly on

[12]This is impossible to verify, however, as it would require a national-level data set which would fully capture the movie release/reel production decisions made by distributors

movie exhibition.[13] The data set is an unbalanced panel across movie theaters and covers the period 2006 - 2011. Attendance figures are broken down by movie, date, screening time and ticket type (child/adult/senior).

Insert Table 1 Here

Contract Data Data on contracts was made available by two of the movie theaters in the sample, henceforth referred to as MT1 and MT2, both of which employ digital projection technology. The data contains a weekly breakdown of revenue split values for each movie screened. The no screen-sharing restraint applies to all movies, while the minimum exhibition period is either two or three weeks for all movies released on the break.

Additional data sources Nationwide viewership trends are sourced from the annual *Theatrical Market Statistics* report release by the Motion Pictures Association of America (MPAA). Movie characteristics and consumer ratings were collected from The Internet Movie Database (IMDB), while professional critic ratings were sourced from Metacritic. Exhibitors provided information on ticket prices. The distribution of consumer demographics was obtained by sampling individuals from the U.S. Census of Population.

Ticket prices The industry practice is to charge the same ticket price irrespective of the movie playing or how long it's been on screen. Given the complex reasons for why uniform pricing is used in the movie exhibition industry (Orbach and Einav, 2007) movie ticket price changes are beyond the scope of this paper. For a given movie theater ticket prices change little over time (no more than once a year in the sample). Additionally, most exhibitors engage in price discrimination between age groups (child/senior discounts), time of day (matinee discounts) and 2D/3D screenings of the same movie.

Insert Table 2 Here

[13] Davis (2006*b*) finds that competition is localized to within a 15 mile radius around a movie theater

4 The Model

This section describes in detail the structural model which aims to explain consumers' demand for movies, exhibitors' scheduling choices and distributors' movie rental pricing. The estimated model is then used to determine the welfare impact of removing the contractual restraints imposed by distributors.

4.1 Demand: Moviegoing Decisions

Market Definition A market is defined to be a movie theater/week combination. Focusing on movie theaters which are local monopolies allows the model to abstract from competitive considerations. The population of the town in which the movie theater is located defines the number of potential customers. The temporal bound of the market should reflect the timeframe over which consumers make decisions and over which they explicitly compare alternatives. Since in the exhibition industry it has become customary for movie theaters to announce their schedules one week at a time, this provides a natural bound in the temporal dimension.

Consumer Decision A discrete choice logit model is used to explain consumers' moviegoing decisions. A consumer is assumed to see a particular movie no more than once.[14] Each week she chooses from among movie/screening time combinations playing at the local movie theater for movies not yet seen.

Agent's Utility Function Consider a movie/screening time combination ms offered by movie theater c in week t, for example Avatar screening on Friday, 9:10pm in movie theater A the week of April 20-26 2012. The indirect utility for consumer i is the form:

$$u_{imsct} = u^M_{imt} + u^S_{imsct} + \Xi_{msct} + \epsilon_{imsct} \tag{1}$$

[14] While this may overestimate the quality of movies such as "Avatar" or "Titanic" which attracted a lot of repeated viewers, such movies are the fringe in the data set and thus any impact this could have on estimation is small

where u^M_{imt} and u^S_{imsct} capture the attractiveness of movie m and screening time s, respectively, to agent i, Ξ_{msct} reflects unobserved utility that is common to all agents and ϵ_{imsct} captures idiosyncratic unobserved utility.[15] If the consumer chooses not to see a movie in market (c,t) she gets $u_{i0ct} = \epsilon_{i0ct}$, where ϵ_{i0ct} is drawn from the same distribution as ϵ_{imsct}.

Movie m's attractiveness to agent i is modeled as follows:

$$u^M_{imt} = \boldsymbol{x}^M_m \boldsymbol{\beta}^M_i + \omega_{im} + I(w_{mt} = 0)\beta^W_1 + w_{mt}\beta^W_2 \tag{2}$$

where $\boldsymbol{x}^M_m$ is a vector of observable movie characteristics, ω_{im} is a consumer-specific, mean-zero, time-invariant signal of whether movie m appeals to him more or less than another consumer with the same characteristics, and w_{mt} is the number of weeks since a movie's nationwide release (thus if a movie is released on the break $w_{mt} = 0$). Together $\boldsymbol{\beta}^M_i$ and ω_{im} capture heterogeneity in consumers' movie tastes.

The attractiveness of screening time s is modeled as follows:

$$u^S_{imsct} = p^A_{imsct}\beta^P + \boldsymbol{I}^S(s)\boldsymbol{\beta}^S_i + x^{3D}_{msct}\beta^{3D}_c \tag{3}$$

where p^A_{imsct} is the price of admission for individual i to screening s, $\boldsymbol{I}^S(s)$ captures which of the screening periods p (e.g. weekday 5-8pm) screening time s falls within, and x^{3D}_{msct} captures whether a specific screening is in 3D.

Heterogeneity in consumer tastes $\boldsymbol{\beta}^M_i$ and β^S_i are consumer-specific coefficients which reflect heterogeneity in moviegoing tastes within the population. They are modeled as multivariate normal with the mean dependent on observable demographic variables and parameters to be estimated, and a variance-covariance matrix to be estimated:

$$\begin{pmatrix} \boldsymbol{\beta}^M_i \\ \boldsymbol{\beta}^S_i \end{pmatrix} = \begin{pmatrix} \boldsymbol{\beta}^M \\ \boldsymbol{\beta}^S \end{pmatrix} + \boldsymbol{\Pi}\boldsymbol{D}_i + \boldsymbol{\Sigma}v_i, \qquad v_i \sim N(0, I) \tag{4}$$

[15] Additive separability between u^M_{imt} and u^S_{isct} allows consumer tastes for movies and screenings to not be correlated. An advantage of this approach over a nested-logit model is that it can capture consumers who prefer movie A to B, but because A does not screen at a convenient time for them they instead choose to see movie B at a more convenient time.

where $\mathbf{\Pi}$ captures how consumer demographics $\boldsymbol{D}_i$ impact their preferences and $\mathbf{\Sigma}$ captures idiosyncratic parameter variance between individuals.[16]

Unobserved product characteristic Ξ_{msct} can be broken down as follows to take advantage of the panel nature of the data set:

$$\Xi_{msct} = \mu_{sct} + \mu_m + \mu_w + \mu_y + \xi_{msct} \tag{5}$$

where all μ parameters can be captured by fixed effects in the estimation stage. The term μ_{sct} captures the attractiveness of screening time s that is specific to movie theater c; it also differs over time t so as to reflect ticket price changes at c. In estimation, the screening period/movie theater/time period fixed effect which captures μ_{sct} also subsumes the price coefficient $p^A_{imsct}\beta^P$ from (3). The term μ_m captures the unobserved quality of movie m; μ_w represents the attractiveness of going to the movies in week $w(t)$, out of 52 weeks total in a year, and thus captures seasonality in the industry, as described by Einav (2007); μ_y captures the annual time trend in attractiveness of going to the movies (relative to, for example, seeing it on DVD/BlueRay as the release window shrinks). ξ_{msct} captures remaining unobserved preferences.

Movie quality uncertainty Combining (1), (2), (3), (4) and (5) yields:

$$u_{imsct} = \delta_{msct} + \omega_{im} + (\boldsymbol{x}^M_m + \boldsymbol{I}^S(s))(\mathbf{\Pi}\boldsymbol{D}_i + \mathbf{\Sigma}v_i) + \epsilon_{imsct} \tag{6}$$

where δ_{msct} is the mean utility of movie/screening time combination $msct$:

$$\delta_{msct} = \underbrace{\boldsymbol{x}^M_m\boldsymbol{\beta}^M + \mu_m}_{\lambda_m} + I(w_{mt} = 0)\beta^W_1 + ... + \xi_{msct} \tag{7}$$

The term λ_m captures the quality of movie m that is not known *ex ante* to the exhibitor - this will feed into the learning model described in section 3.2.

Market shares Given the choice model described above, the set of consumers who choose combination (m, s) in market (c, t) is defined as

$$\begin{aligned}&\mathcal{A}_{msct}(\boldsymbol{x}_{msct}, \mathbf{\Xi}_{msct}; \boldsymbol{\theta}) = \\ &\quad \{(\boldsymbol{D}_i, \epsilon_{imsct}) | u_{imsct} > u_{im's'ct} \forall m', s' s.t. (m', s') \neq (m, s) \text{ and } m \notin \iota_{it}\}\end{aligned} \tag{8}$$

[16]The empirical implementation uses a diagonal $\mathbf{\Sigma}$, although correlations between coefficients can easily be added

where $\boldsymbol{x}_{msct}$ and $\boldsymbol{\Xi}_{msct}$ are the observable and unobservable characteristics, respectively, of combination (m, s), while $\boldsymbol{\theta}$ includes all model parameters.

Modeling Moviegoing History One innovation in the context of literature on movie exhibition industry is that the model explicitly keeps track of movies seen by consumer i up to period t, ι_{it} and ensures consumers see a given movie only once. This allows for explicit modeling of consumer selection, wherein consumers who choose to see the movie in the early weeks are likely to have higher valuation for this movie compared to those who see it in later weeks. Previous work, such as Einav (2007), estimated a "decay factor" which captured in a reduced-form way how a movie's box office draw falls as time passes since its nationwide release date. The approach adopted in this model disentangles the "decay factor" into three components: (1) with each week since a movie opens at a movie theater more people see it, shrinking the potential market (2) people who value a movie highly are more likely to see it early, and thus the average consumer who has not yet seen the movie in later weeks is likely to value it less and is thus less likely to see it (3) a concentrated advertising campaign gives people higher utility from seeing a movie early.

4.2 Supply: Exhibitor Scheduling Problem

The exhibitor aims to maximize the present discounted value of expected profits $\mathbb{E}[\sum_{\tau=t}^{\infty} \beta^{\tau-t} \pi_{\tau}^{*} | \boldsymbol{s}_{ct}, \boldsymbol{r}_{ct}^{*}, \mathcal{M}_{ct}^{*}]$ by choosing which movies to screen each period: $\{\mathcal{M}_{c\tau}\}_{\tau=t}^{\infty}$. This leads to the Bellman equation:

$$V(\boldsymbol{s}_{ct}, \boldsymbol{r}_{ct}^{*}, \mathcal{M}_{ct}^{*}) = \max_{\mathcal{M}_{ct} \subset \mathcal{M}_{ct}^{*}} \pi^{*}(\boldsymbol{s}_{ct}, \boldsymbol{r}_{ct}^{*}, \mathcal{M}_{ct}) + \beta E\left[V(\int_{ct+1}, \nabla_{ct+1}^{*}, \mathcal{M}_{ct+1}^{*}) \middle| \boldsymbol{s}_{ct}, \boldsymbol{r}_{ct}^{*}, \mathcal{M}_{ct}\right]$$

where $\boldsymbol{s}_{ct}$ are the state variables, $\mathcal{M}_{ct}^{*}$ are the movies offered by distributors to movie theater c in period t, $\boldsymbol{r}_{ct}^{*}$ are revenue splits set by the distributors, π^{*} is the per-period expected profit function and β is the discount factor. Since it is close to costless for distributors to make all their movies available to exhibitors employing digital projection technology, such as MT1 and MT2, $\mathcal{M}_{ct}^{*}$ is equal to the set of all movies released nationwide in period t.

In each period t, given a set of movies $\mathcal{M}_{ct}$ to screen, the exhibitor decides when to screen them. His per-period profit given schedule $\boldsymbol{z}_{ct}$ is:

$$\pi(\boldsymbol{s}_{ct}, \boldsymbol{r}^*_{ct}, \boldsymbol{z}_{ct}) = \sum I_{msct}(\boldsymbol{z}_{ct}) \int_{\mathcal{A}_{msct}} (p^A_{imsct}(1 - r^*_{mct}) + \pi^C_c) dP^*(\epsilon) dP^*(D) - C^C_c \tag{9}$$

where $I_{msct}(\boldsymbol{z}_{ct})$ is a an indicator function whether movie/screening combination ms is part of schedule $\boldsymbol{z}_{ct}$, π^C_c is the average concession profits per moviegoer in movie theater c, while C^C_c is the fixed cost of keeping movie theater c open for one period.[17] Maximizing the profit function produces the optimal schedule for the set of movies $\mathcal{M}_{ct}$:

$$z^*(\boldsymbol{s}_{ct}, \boldsymbol{r}^*_{ct}, \mathcal{M}_{ct}) = \underset{\boldsymbol{z}_{ct} \in Z(\mathcal{M}_{ct})}{\arg\max}\ \pi(\boldsymbol{s}_{ct}, \boldsymbol{r}^*_{ct}, \boldsymbol{z}_{ct}) \tag{10}$$

where $Z(\mathcal{M}_{ct})$ is the set of all possible schedules for exhibitor c in period t given $\mathcal{M}_{ct}$. The per-period expected profit function used in (9) is then:

$$\pi^*(\boldsymbol{s}_{ct}, \boldsymbol{r}^*_{ct}, \mathcal{M}_{ct}) = \pi(\boldsymbol{s}_{ct}, \boldsymbol{r}^*_{ct}, z^*(\boldsymbol{s}_{ct}, \boldsymbol{r}^*_{ct}, \mathcal{M}_{ct})) \tag{11}$$

Contractual restraints The no screen-sharing restraint enters the exhibitor scheduling problem through the function $Z(\cdot)$ - if screen-sharing is allowed the movie theater has more flexibility with its schedule, and thus

$$Z(\mathcal{M}_{ct})_{\text{no screen-sharing}} \subset Z(\mathcal{M}_{ct})_{\text{screen-sharing}}$$

The minimum exhibition period restraint enters the exhibitor scheduling problem through $\mathcal{M}_{ct}$ e.g. if exhibitor c commits in period $t-1$ to screening movie m' for 2 periods, then every combination $\mathcal{M}_{ct}$ in period t will contain m'. Thus

$$\{\mathcal{M}_{ct}\}_{\text{minimum exhibition period}} \subset \{\mathcal{M}_{ct}\}_{\text{no minimum exhibition period}}$$

[17] Per person concession sales are assumed to be constant in the data set, which is supported by Gil and Hartmann (2007) who find that concession sales are roughly proportional to total attendance in Spanish movie theaters. Fixed costs are independent of the exhibitor's choice of movies played and the schedule, which is a reasonable assumption as long as the opening hours/days and number of screens operating are kept constant.

Ex ante movie quality signals and learning Before a movie is released nationwide each exhibitor receive a signal of its quality that is specific to him:

$$\hat{\lambda}_{mc} = \lambda_m + \nu_{mc} \qquad\qquad \nu_{mc} \sim N(0, \sigma^2_{\nu c}) \tag{12}$$

which depends on the true movie quality, λ_m, and an idiosyncratic term ν_{mc}.

The exhibitor learns a movie's true quality λ_m one week after the movie's nationwide release. This perfect learning assumption reflects the fact that the exhibitor can learn a movie's true quality even if he does not screen it, by analyzing widely-available box office revenue information which captures how well the movie did in its first week of release.[18]

State variables State variables include, for week t:

- $\{\iota_{it}\}_{\forall i}$: the set of movies seen by consumer i up to period t;
- $\{\hat{\lambda}_{mc}\}_{\forall m \in \mathcal{M}^*_{ct}}$: exhibitor c's belief about unknown movie quality for movies he's considering playing in week t;

4.3 Supply: Distributor Decision

Each period the distributor chooses the revenue splits, $\boldsymbol{r}^*_{ct}$ at which to offer his movies to exhibitor c. Following Ho (2009), these rental prices are determined by a take-it-or-leave-it bargaining model. Distributors negotiate terms for each of their movies separately, while exhibitors negotiate separately for each movie theater they own. Negotiations take place before period t for all (c,m) combinations where $m \in \mathcal{M}_{-c}$, the set of movies not already screened by c.

Let $\boldsymbol{r} = \{\boldsymbol{r}_{mc}\}_{\forall m,c}$ be the set of revenue splits for all movie theaters and movies, where the revenue split time profile $\boldsymbol{r}_{mc}$ maximizes the profits movie

[18] Moreover, by assuming learning happens independently of exhibitor's actions the estimation procedure is reduced to a static problem. An alternative approach explicitly modeling learning in a bayesian fashion, following Hitsch (2006), was deemed inappropriate in this setting because it cannot reconcile learning from own's experience and from other sources

m's distributor derives from screening it at movie theater c:

$$\Pi_{mc}(\boldsymbol{r}_{mc}, \boldsymbol{r}_{-mc}, \boldsymbol{s}_{ct}) = \mathbb{E}\Big[\sum_{t=1}^{\infty} \beta^t \sum I_{msct}\big(z^*(\cdot)\big) \int_{\mathcal{A}_{msct}} p^A_{imsct} r_{mct} dP^*(\epsilon) dP^*(D)\Big] \tag{13}$$

where $z^*(\cdot)$ is the optimal schedule movie theater c arrived given its own movie quality signals, while $\mathcal{A}_{msct}$ is correspondingly calculated using true movie quality for movies played in previous periods, and the ex ante quality signals otherwise.[19] If an agreement is not reached between movie theater c and movie m the set of revenue splits for all other movie theater/movie pairs, $\boldsymbol{r}_{-me}$, is not renegotiated. The resulting bargaining equilibrium is:

$$\boldsymbol{r}^*_{mc} = \arg\max_{\boldsymbol{r}_{mc}} \Pi_{mc}(\boldsymbol{r}_{mc}, \boldsymbol{r}_{-mc}, \boldsymbol{s}_{ct}) \tag{14}$$

5 Estimation, Identification and Counterfactual Calculation

Estimation proceeds in two stages: section 4.1 describes how demand model parameters are estimated using the BLP approach augmented with micro-moments à la Petrin (2002), while section 4.2 describes how observed schedules and revenue splits are used to estimate parameters driving the movie quality generation process. The algorithm used to calculate exhibitor and distributor decisions in the counterfactuals are described in the Appendix.

5.1 Demand Model Estimation

Primary moments The estimation strategy follows the standard GMM approach established by BLP. By assuming that ϵ_{imsct} is distributed i.i.d. Type I extreme value (Berry, 1994) an inversion produces the residual $\xi_{msct}(\boldsymbol{\theta})$ which

[19] C^M_{mct}, the per-movie cost a distributor needs to incur to provide a copy of movie m to movie theater c in period t, is close to zero for MT1 and MT2 and thus omitted

is then used to construct the primary set of moments:

$$E[\boldsymbol{G}_1(\boldsymbol{\theta})] = E[\boldsymbol{Z}_{msct} \cdot \xi_{msct}(\boldsymbol{\theta})] \tag{15}$$

where $\boldsymbol{Z}_{msct}$ is a vector of instruments that are orthogonal to ξ_{msct}.

Instruments and price endogeneity Movie ticket prices do not vary between movies - all price variation is driven by price discrimination (see Table 2). The price coefficient can thus be decomposed as follows:

$$p^A_{imsct} = p^A_{sct} + p^A_{ic} + p^{3D}_c x^{3D}_{msct} \tag{16}$$

where p^A_{sct} captures the "base" price, matinee discounts and price changes over time, p^A_{ic} is the child/senior discount and p^{3D}_c is the premium for a 3D screening; all three are exhibitor-specific. In estimation, all three components are captured using fixed-effects, thus ensuring they do not enter $\xi_{msct}(\boldsymbol{\theta})$. This removes the price endogeneity that models using the BLP framework usually have to take into account, and . As a result, all independent observable variables in (7) and fixed-effects are exogenous and thus are valid instruments in constructing the primary moments.

Micro Moments The estimation procedure is augmented by five sets of additional micro-moments, following Petrin (2002). The first three sets are derived from information in the MPAA 2010 Theatrical Market Statistics and allow for better identification of heterogeneity by age within the population. They include 1. absolute moviegoing frequency 2. relative moviegoing frequency by age group 3. age composition of the frequent moviegoer group. The forth and fifth sets are derived straight from the data: 4. moviegoing frequency by movie theater, and 5. attendance at each screening by age group. The latter takes advantage of child/adult/senior ticket breakdown which is available for some but not all movie theaters in the data. In total there are 22 micro moments: $E[\boldsymbol{G}_2(\boldsymbol{\theta})]$

The Objective Function The two sets of moments that enter the GMM objective function are $\boldsymbol{G}_1(\boldsymbol{\theta})$, the standard BLP moments, and $\boldsymbol{G}_2(\boldsymbol{\theta})$, the micro moments. The population moment conditions are assumed to uniquely equal zero at the true $\boldsymbol{\theta}_0$:

$$E[G(\boldsymbol{\theta}_0)] = E\begin{bmatrix}\boldsymbol{G}_1(\boldsymbol{\theta}_0)\\ \boldsymbol{G}_2(\boldsymbol{\theta}_0)\end{bmatrix} = 0 \tag{17}$$

The GMM estimator then takes the form

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} G(\boldsymbol{\theta})' W^{-1} G(\boldsymbol{\theta}) \tag{18}$$

where W is a weighting matrix set to be $Z'Z$. In order to estimate standard errors the approach developed by Hansen (1982) is followed, which allows both sampling error and simulation error to be taken taken into account. Standard errors are clustered - see Berry, Levinsohn and Pakes (2004) for details.

Compensating variation calculations Because the coefficient on ticket price, β^P, is not separately identified in the model Compensating Variation (CV) as developed by Hicks (1939) cannot be used to measure the change in consumer welfare. Instead, a quasi-CV measure is defined as the monetary value of movie/screening combinations a consumer would have to see to make up for screenings lost due to the imposition of vertical restraints.

5.2 Estimation of Movie Quality Generating Process

A big challenge of the estimation procedure is to back out true movie quality values and *ex ante* quality signals for movies observed in the data set, and to construct a way to simulate them for movies not in the data set. The following equation summarizes the relationship between predicted movie quality ($x_m^M \boldsymbol{\beta}^M$), true movie quality ($\lambda_m$) and *ex ante* movie quality signals ($\hat{\lambda}_{mc}$):

$$\hat{\lambda}_{mc} = \underbrace{x_m^M \boldsymbol{\beta}^M + \mu_m}_{\lambda_m} + \nu_{mc} \qquad \mu_m \sim N(0, \sigma_\mu^2) \qquad \nu_{mc} \sim N(0, \sigma_{\nu c}^2) \tag{19}$$

The demand model estimates λ_m for movies in the data set, however estimating $\boldsymbol{\beta}^M$ is more complicated than simply regressing λ_m values on $\boldsymbol{x}_m^M$ for these movies. Since these movies were selected for screening by the exhibitor because he knew or expected them to attract higher audiences than movies he chose not to screen, they most likely exhibit, on average, positive ν_{mc} and μ_m values. A simple OLS regression would thus result in biased estimates. Instead, a model is necessary which will fully capture the movie quality generation and the exhibitor movie selection processes, which will also allow for the estimation of movie quality signals $\hat{\lambda}_{mc}$.

This section describes the two-stage procedure used to identify $\hat{\lambda}_{mc}$ for movies in the data set, bounds on $\hat{\lambda}_{mc}$ and λ_m values for movies not observed in the data set, as well as parameters in the movie quality generating process $(\boldsymbol{\beta}^M, \sigma^2_\mu, \sigma^2_{\nu c})$. The first stage sets up bounds on movie quality signals based on revealed schedule and pricing decisions in order to identify signals for movies screened *relative* to the quality of the best alternative. The second stage uses a maximum likelihood estimator to identify the *absolute* value of the exhibitor's movie quality signals, as well as parameters in (19). Due to lack of revenue split information for three of the exhibitors, the following estimation is only carried out for $c \in \{\text{MT1,MT2}\}$.

5.2.1 Stage 1: Movie quality bounds

Consider movies opening at movie theater c in period t: $\mathcal{M}^+_{ct}$. For those which are released on the break, $\mathcal{M}^+_{ct}(w_{mt} = 0)$, the exhibitor does not know λ_m and bases his decision on his *ex ante* signals $\hat{\lambda}_{mc}$. Knowing the exhibitor's signals the distributor chooses a price $\boldsymbol{r}_{mc}$. These two decisions help establish bounds on $\hat{\lambda}_{mc} \forall m \in \mathcal{M}^+_{ct}$:

1. Lower bound $\underline{\hat{\lambda}_{mc}}$: movie m is at least as good as the best alternative from the set of movies available to the exhibitor: $\mathcal{M}^*_{ct}$ (otherwise he would screen one of those movies instead)

2. Upper bound $\overline{\hat{\lambda}_{mc}}$: movie m is not good enough that the distributor could increase $\boldsymbol{r}_{mc}$ and the exhibitor would still choose to screen the movie[20]

[20] $\boldsymbol{r}_{mc}$ takes on discrete values, otherwise the upper bound would be equal to the lower bound. Also, this bound is nonexistent if $\boldsymbol{r}_{mc}$ is already at the highest value it can take (for

These bounds are expressed relative to the quality of the best alternative available to exhibitor c in period t: $\hat{\lambda}_{0ct}$. Which exact movie constitutes the best alternative is not known to the econometrician - it could either be a movie released this period ($w_{mt} = 0$), in which case $\hat{\lambda}_{0ct}$ will equal the *ex ante* quality signal for this movie, or a movie released in a previous period ($w_{mt} > 0$), in which case $\hat{\lambda}_{0ct}$ will equal the true movie quality of that movie:

$$\hat{\lambda}_{0ct} = \max \left[\max_{m \in \mathcal{M}_{ct}^{-}(w_{mt}=0)} \left(\hat{\lambda}_{mc} \right); \max_{m \in \mathcal{M}_{c}^{-}(w_{mt}>0)} \left(\lambda_m + w_{mt} \beta_2^W \right) \right] \tag{20}$$

where $\mathcal{M}_{ct}^{-}$ is the set of movies not released in period t. Three cases are possible for each period:

1. $\mathcal{M}_{ct}^{+} = \emptyset$: no movies opened in period t
2. $\mathcal{M}_{ct}^{+} = \mathcal{M}_{ct}^{+}(w_{mt} = 0)$: all movies that opened in period t were also released in period t
3. $\mathcal{M}_{ct}^{+}(w_{mt} > 0) \neq 0$: at least one movie that opened in period t was released before period t

In case 1, not enough information is available regarding the *ex ante* quality signals of movies which were released this period. That none of them were picked up by the exhibitor may be because they are all of poor quality, but it may also be because no screens could be freed up for a new movie because of minimum exhibition period restrictions on movies already being screened. In case 2 there is not enough information to identify the absolute $\hat{\lambda}_{0ct}$ value at this stage, and bounds on $\hat{\lambda}_{mt}$ are calculated relative to a range of possible $\hat{\lambda}_{0ct}$ values. In case 3, the absolute value of $\hat{\lambda}_{0ct}$ can be set-identified relative to $\{\lambda_m\}_{m \in \mathcal{M}_{ct}^{+}(w_{mt}>0)}$ ie. the fact that a movie of known quality λ_m was released provides a lower and upper bound on what the best alternative available to the exhibitor in time period t was. In case 3 absolute bounds on $\hat{\lambda}_{mt}$ can thus be calculated at this stage; in cases 1 and 2 their absolute values can only be calculated in stage 2 of the estimation procedure.

a full discussion on discretization and bounding of $\boldsymbol{r}_{mc}$ values see Appendix)

5.2.2 Stage 2: Maximum likelihood estimation

Building on relative bounds for $\hat{\lambda}_{mc}$ and $\hat{\lambda}_{0ct}$ estimated in Stage 1, the second stage uses maximum likelihood estimation to identify the absolute values of these bounds, as well as estimate parameters $\boldsymbol{\beta}^M$, σ_μ^2 and $\{\sigma_{\nu c}^2\}_{\forall c}$ in (19). Intuitively, the log-likelihood function is used to reconcile movie quality bounds established in Stage 1 with estimates of movie quality λ_m (for movies in the data set) and movie observable characteristics x_m^M (for movies not in the data set). For full specification of the log-likelihood function $\ell(\cdot|\cdot)$ see 7.3.3.

Computation The estimation procedure has two loops. The outside loop is a non-linear search over possible $\boldsymbol{\beta}^M, \sigma_\mu^2$ and $\{\sigma_{\nu c}^2\}_{\forall c}$ values. Given a multiple $\{\boldsymbol{\beta}^M, \sigma_\mu^2, \{\sigma_{\nu c}^2\}_{\forall c}\}$ of candidate values the inner loop chooses a set of $\{\hat{\lambda}_{0ct}, \lambda_{0t}\}_{\forall c, t \in T_c}$ values which maximize the log-likelihood function:

$$\ell(\sigma_\mu^2, \{\sigma_{\nu c}^2\}_{\forall c}, \boldsymbol{\beta}^M|\cdot) = \max_{\{\hat{\lambda}_{0ct}, \lambda_{0t}\}_{\forall c, t \in T_c} \in \boldsymbol{\lambda_0}} \ell(\sigma_\mu^2, \{\sigma_{\nu c}^2\}_{\forall c}, \boldsymbol{\beta}^M, \{\hat{\lambda}_{0ct}, \lambda_{0t}\}_{\forall c, t \in T_c}|\cdot) \tag{21}$$

Intuitively, identification of $\boldsymbol{\beta}^M$ and σ_μ^2 comes from analyzing λ_m values for movies screened by the exhibitors as well as λ_{0t} value for movies not screened. Identification of $\sigma_{\nu c}^2$ comes from comparing bounds on $\hat{\lambda}_{mc}$ to λ_m for movies screened and $\hat{\lambda}_{0ct}$ to $\boldsymbol{x}_m^M \boldsymbol{\beta}^M$ for movies not screened (the latter also help to identify σ_μ^2). The estimation procedure is sufficiently quick that confidence intervals can be computed using a standard bootstrap, resampling at the time period level. A detailed description of the estimation algorithm can be found in the Appendix.

Implementation Up to this point the exact metric exhibitors use when determining whether one movie is "better" than another has purposefully been left unspecified. Implementing a fully dynamic estimation process with forward-looking exhibitors, as described in section 3.2, has proven computationally intractable.[21] To make the estimation procedure feasible a simplify-

[21] Unlike most settings in which the estimation procedure is fully dynamic, exhibitors' choice set changes over time which means the value-function is non-stationary

ing assumption is made that when making scheduling and pricing decisions exhibitors and distributors only consider profits in the current period, rather than over the whole period each movie is expected to be screened. This assumption is supported by conversations with exhibitors, who say that in the face of uncertainty about movies' true quality they focus on first-week profits when making scheduling decisions, and only consider binding restraints once true movie quality is revealed in later periods.

Nonetheless, assuming exhibitor and distributor myopia can lead to estimation bias. Specifically, if exhibitors' ability to release blockbusters on the break is inhibited by minimum exhibition period restraints on other movies being screened, it is possible they incorporate releases of such blockbusters when making their scheduling decisions. If this is in fact the case one should expect exhibitors to take on fewer movies in the period leading up to the release of a big blockbuster. Whether they do this can be tested empirically. Table 3 reports, for exhibitors in the sample, the average proportion of new movies compared with all movies being screened in a given week. Strategic consideration for three factors is analyzed: 1. release of blockbuster movies 2. high-attendance weeks 3. holidays.[22] The results show that exhibitors take on no fewer new movies right before blockbuster movies are released than they do on average - this should alleviate the concern that they act strategically in a way not captured by the model.

Insert Table 3 Here

Moreover, any potential bias resulting from not accounting for exhibitors' forward-looking behavior would lead to a reduction in expected improvements from removing vertical restraints. If exhibitors are in fact less likely to take on new movies right before blockbuster movies are released, the current model assuming myopia would bias $\hat{\lambda}_{0ct}$ and λ_{0t} estimates downward (as the exhibitor

[22] Blockbuster movies were identified by the author, while high-attendance weeks were defined as weeks with cumulative attendance higher than 120% of the average weekly attendance

not taking on some of the outside movies would be “blamed” on these movies’ poor quality). Such a bias would reduce the simulated movie quality values for movies not in the data set, thus depressing gains in welfare from lifting the restraints (when exhibitors are likely to take on some of these movies).

6 Results

The structural model is estimated and the results are presented below. The section that follows uses the estimated model to conduct counterfactual simulations of what would be the effect of removing the contractual restraints; this is done for the two movie theaters for which revenue split information is available. Also, it analyzes the payoffs faced by distributors when deciding whether to impose restraints.

6.1 Parameter Estimates

Insert Table 4 Here

The demand parameter estimates presented in Table 4 are intuitive and coefficients have the expected directionality. Consumer differ significantly in their desire to go to the movie theater, and young people aged 12-24 have an especially high valuation of going to the movies. On net, the average moviegoer prefers to see a movie in 3D despite its higher price. Screening a movie on the break provides a big boost in attractiveness, while the longer it is been since a movie’s nationwide release the less attractive it is. The fall in attractiveness is relatively slow, however, suggesting most of the decline of a movie’s box office revenue take in the weeks following its release is driven by a shrinking pool of people interested in seeing it.

Age proves to be a strong predictor of consumers’ tastes in movies. As expected, PG-13 and R-rated movies are considered less appealing to young people aged 2-11 by those who take them to the movie theater, while family and animated movies are considered most appealing. People aged 12-59 are less likely to go to movies before 5pm during the week because of school/work commitments. Finally, consumers differ in their preferences for individual movies, as shown in the high standard deviation of ω_{im}. This helps explain why blockbusters such as *Avatar* never grab the whole market to themselves.

Insert Figure 1 and Figure 2 Here

Figure 1 plots week fixed effects. It shows that going to the movies is more attractive during and around major holidays, which corresponds to findings in Einav (2007), though late summer months are found to be less attractive.

Figure 2 plots normalized time period fixed effects, by movie theater. As expected, the normalized fixed effects exhibit similar trends across different movie theaters. For all but one movie theater the *Mon-Thu, after 8pm* time period is the least attractive to moviegoers, while the *Friday, 5-8pm* and *Sat/Sun, 5-8pm* periods are the first and second most attractive overall.

Insert Table 5 Here

Table 5 shows the top and bottom 10 movie fixed effects, alongside each movie's national Box Office Revenue take. The first take-away is that the difference in fixed effects value between the top and bottom movie observed in the sample is considerably larger in magnitude than that for either week- or time period-fixed effects. This emphasizes the importance to exhibitors of choosing the best movies each week, and suggests that allowing exhibitors to drop poorly performing movies quickly by lifting the minimum exhibition period will allow them to substantially boost attendance. The movies in the top 10 were some of the highest grossing nationwide over the sample period, while the bottom 10 movies all did poorly in their theatrical runs.

Insert Table 6 Here

Table 6 reports results from the second stage estimation process, which identifies parameters driving the movie quality generating process (19). Movies that have better "word of mouth", as proxied by *IMDB moviegoer rating*, have higher appeal for moviegoers.[23] Movies which receive a G-rating from the MPAA, meaning they are suitable for all audiences, have the highest average appeal. As expected, movies with a higher budget are, on average, more attractive to moviegoers. This captures two factors: (1) such movies are in fact better (2) a higher budget proxies for more promotional activity. Although only some of the *genre* coefficients turned out to be significant, it is clear that musicals and documentaries hold less appeal for audiences, while fantasy and sci-fi movies are more attractive *ceteris paribus*. The estimated σ_μ coefficient is larger than either of the $\sigma_{\nu c}$ coefficients - this indicates exhibitors' signals of movie quality are relatively accurate compared to how much true quality varies across movies with the same observable characteristics x^M. It also underlines the importance of explicitly estimating the *ex ante* movie quality signal exhibitors receive - the signal is a much better predictor of true movie quality than simply calculating $\boldsymbol{x}_m^M \boldsymbol{\beta}^M$ based on movies' observable characteristics.

6.2 Counterfactual Results: Welfare Cost of Restraints

Table 7 presents results of counterfactual simulations which measure the impact of removing the contractual restraints. The simulations were carried out for the two movie theaters for which revenue split data was available, MT1 and MT2. Since the minimum exhibition period is not observed in the sample it is impossible to set up a single base case simulation which accurately reflects the actual restraints. Instead, the base case is split into two scenarios.

Insert Table 7 Here

[23] Although a movie's IMDB rating can change over time, anecdotal evidence suggests the changes are small and their direction is random, which ensures the variable is a good proxy for "word of mouth". For an example of how "word of mouth" can be modeled explicitly in the movie exhibition industry see Moretti (2011).

In scenario 1. the minimum exhibition period restraint imposed on all movies released on the break is the minimum of three weeks or the number of weeks a movie was actually screened for.[24] In scenario 2. the minimum exhibition period period is two weeks. In reality the minimum exhibition period restraints lie between these two scenarios. Scenario 3. lifts the minimum exhibition period restraint but still disallows screen-sharing. Scenario 4. lifts both restraints.[25] Comparing scenarios 3. and 4. to base cases 1. and 2. provides upper and lower bounds, respectively, on the improvements expected from removing the restraints.[26] For every scenario multiple simulation were performed, each based on a different set of randomly simulated $\hat{\lambda}_{mc}$ and λ_m values, and the final numbers presented are an average across all simulations.

Change in attendance Removing the minimum exhibition period restraint only (scenario 3.) results in a substantial increase in attendance compared to the base cases: 3.8 to 8.7% for MT1 and 6.8 to 12.8% for MT2. As expected, these values are greater for the smaller movie theater MT2 where the constraints are more restrictive. One factor contributing to the rise in attendance is the increase in the number of movies screened. The increase is considerably higher for MT2, supporting the notion that smaller movie theaters have more to gain from offering a wider choice of movies to potential moviegoers. The rise in attendance is also driven by the fact that once the minimum exhibition period restraint is lifted exhibitors are able to quickly adjust their schedules once they learn the true movie quality.

[24] A four week minimum, vary rare according to exhibitors, was not implemented because even if it was imposed it would be attached to blockbusters for which it would not bind

[25] Allowing for screen sharing while retaining the minimum exhibition period restraint would in effect invalidate the latter restraint, allowing the exhibitor to put one screening of a movie per week to technically satisfy that restraint even though it would violate its intent

[26] Ideally, scenarios 3. and 4. could be compared directly to the observed schedules and attendance figures. Unfortunately, the complexity of the demand and supply models employed means that while observed attendance lies between that predicted in scenarios 1. and 2., other metrics are less well aligned. Comparing simulated figures from scenarios 3. and 4. to those from scenarios 1. and 2. provides a more meaningful comparison.

Allowing for screen-sharing (scenario 4.) on top of removing the minimum exhibition period restraint results in a further increase in attendance and number of movies screened, although these gains are relatively smaller than those realized in scenario 3.. This suggests the current strategy of showing one movie per screen throughout the week does not leave a lot of untapped demand.

Welfare impact The rise in attendance results not only in substantial welfare increases for consumers, but also in substantial increases in profits for exhibitors and distributors. Consumer welfare rises slightly more than attendance under most scenarios, suggesting people like the movies they see better. Higher profits are not distributed uniformly - exhibitors capture more of the incremental profits than do distributors. This reflects the fact that removing contractual restraints broadens the exhibitors' strategic options and improves their bargaining position.

6.3 Counterfactual Results: Distributors' Incentives for Imposing Restraints

Although results presented in section 5.2 show that lifting vertical restraints would increase total distributor profits for the digital movie theaters in the sample, to date distributors have not relaxed the restraints on these movie theaters. In fact, even though digital projection is now the dominant technology in the industry, there are no signs the restraints may be relaxed nationwide.

One possible explanation as to why distributors continue to impose restraints may be provided by performing a game theoretic analysis of the movie theaters in the sample. The process of imposing/lifting restraints is formulated as a one-period game between the Majors.[27] In the game each distributor decides, given his expected payoffs, whether to impose restraints on exhibitors

[27] Although there were over 50 distributors operating in the U.S. market in 2011, the market is highly concentrated and the top 6 Majors regularly accounted for around 85% of box office revenues in 2011. It thus makes sense to perform the analysis from their perspective, as smaller players are likely to follow them.

screening his movies. Table 8 analyzes the stability of two potential equilibria: 1. no distributors impose restraints, and 2. all distributors impose restraints.

Insert Table 8 Here

The first take-away from Table 8 is that when all distributors impose restraints it is not profitable for any player to deviate and lift restraints for his movies. Consider what happens when Major #1 lifts the restraints on his movies: the exhibitors are now able to more quickly replace the low true quality λ_m movies they had taken on the basis of deceptively high $\hat{\lambda}_{mc}$ signals. Major #1 thus loses out as his movies are replaced, mostly with offerings from his competitors (though some will be replaced with other movies from Major #1). Empirical calculations show that because other distributors continue imposing restraints the overall gain in attendance is not enough to compensate Major #1 for this loss in profits.

The second take-away is that it is not an equilibrium for none of the players to impose restraints, as every Major would then have an incentive to unilaterally deviate. The intuition is the corollary of that presented above - by imposing restraints Major #1 ensures that exhibitors who receive a high $\hat{\lambda}_{mc}$ signals on a movies with low true quality λ_m screen them for a longer period of time. This translates into substantial profit gains for the deviating distributor.

These findings are noteworthy, as they show that vertical restraints may persist even though all parties imposing them would be better off if they were lifted. Because they lack unilateral incentives to lift the restraints distributors find themselves in a prisoner's dilemma game, as described in Besanko and Perry (1993). The mechanism through which this happens is similar in both cases - in that paper when manufacturers adopt restraints they internalize a positive externality they imposed on their competitors; here, distributors impose a negative externality on their competitors. Although it is not possible to provide a definitive answer as to why distributors impose restraints on digital movie theaters without being able to view the whole U.S. movie exhibition market, these findings provide a possible explanation and suggest it would be welfare-improving for competition authorities to ban the use of these restraints for digital movie theaters.

A related question is whether the socially optimal result of no vertical restraints being imposed can be achieved without top-down bans imposed by the antitrust authorities. A possible solution would be to insert a "most-favored nation" clause into the distributor-exhibitor contract. Under such an agreement the distributor would not impose any restraints on its movies as long as none of his competitors did so. This would ensure the stability of the socially optimal "no restraints" equilibrium. From a game theoretic perspective such an agreement corresponds to a grim trigger punishment strategy in an infinitely repeated prisoner's dilemma game. It should be noted that such agreements would be labeled "contracts that reference rivals", a type of contract that is viewed as potentially anti-competitive by the U.S. antitrust authorities. The reason for this is that they can be used to make stable collusive agreements where the participants would otherwise deviate. This ability is used here to maintain a socially optimal equilibrium.

6.4 Lessons for the U.S. Movie Exhibition Market

It is worth considering whether the findings from this small sample of exhibitors have any implications the U.S. movie exhibition market as a whole. Any findings have to be seen through the prism of the way in which the sample is not representative of the whole market: (1) movie theaters are local monopolies, (2) they are all small and medium-sized, and (3) the exhibitors are all independent. These limitations are addressed below.

First, consider what happens when there are multiple movie theaters in one market. Conversations with exhibitors suggest movie theaters in such markets aim to offer a selection of movies which is no worse than their competitors'. This is borne out in the real world, where anecdotal evidence suggest movie theaters in the same city offer very similar schedules. This lack of scheduling complementarity suggests the combined choice set from competing movie theaters is likely to be not much more varied than that offered by an individual movie theater. Combined with the finding in Davis (2006*a*) that there is little business-stealing between movie theaters in one city, this suggests welfare gains in competitive markets are likely to be similar to those for local monopolies.

Second, since larger movie theaters can take on more movies each week it follows the benefits from removing restraints will be smaller. This is supported in the counterfactual experiments, where welfare gains are smaller for the larger of the two movie theaters. Although it is impossible to conclude how much welfare gains from removing contractual restraints are reduced for large movie theaters without access to detailed attendance data for such movie theaters, they will always remain positive. In addition, the types of movie theaters observed in the data set (between 1 and 7 screens) represented 75% of movie theaters[28] in the U.S. in 2000 (Davis, 2006*a*).

Third, conversation with exhibitors suggest movie theaters which are part of a chain (e.g. AMC Loews, Regal Entertainment) do not get preferential treatment when it comes to restraints.[29] While large chains are understood to be able to negotiate better rental prices down for some movies, this determines the split of the increased profits pool rather than its size. Total welfare gains at movie theaters owned by large chains are thus not expected to be different to those independently owned.

7 Conclusion

This paper shows that vertical restraints imposed by distributors on digital movie theaters significantly can reduce consumer welfare and industry profits.

[28]This figure, the only one available, is calculated based on the number of movie theaters; the corresponding figure calculated based on box office revenues should be expected to be lower but still significant

[29]It should be noted that large exhibitors could not be reached for comment; however, it is reasonable to expect that if large chains were thought to receive preferential treatment their smaller competitors would point that out

Removing them would allow exhibitors to screen better movies by quickly adjusting their schedules in response to learning the true quality of movies, and would increase the variety of the consumer choice set by allowing them to screen more movies overall. This finding has implications for the U.S. exhibition industry as it concludes its conversion to digital projection.

This paper also suggests a possible explanation for why restraints persist in this industry even though everyone, including movie distributors who impose them, would be better off if they were lifted. It provides empirical evidence for the finding of Besanko and Perry (1993), who show that vertical restraints can arise as a result of a prisoner's dilemma game. The central question in the literature has been whether vertical restraints are imposed because they have pro-efficiency or anti-competitive impact; the findings from this paper suggest the circumstances under which they are imposed also play a key role.

The analysis conducted in this paper suggests it would be welfare-improving for competition authorities to ban distributors from imposing restraints for digital movie theaters. Alternatively, the contracts between distributors and exhibitors could be modified with a "most-favored nation" clause, which would ensure the stability of the no-restraints equilibrium.

Beyond movie exhibition, this paper's findings confirm that in industries where products or services are sold to consumers through independent retailers non-price vertical restraints imposed by manufacturers can have a significant impact on total welfare. When only a select group of products can be offered to consumers at a given time and place, such restraints can substantially alter the composition of consumers' choice sets and thus their purchasing decisions, even if they do not impact prices. The impact is likely to be greater when retailers regularly make decisions under uncertainty as to the appeal of new products to consumers. This has implications for industries such as radio and TV, offline and online advertising, as well as retail sales.

References

Alimurung, Gendy. 2012. "Movie Studios Are Forcing Hollywood to Abandon 35mm Film." *http://www.laweekly.com/2012-04-12/film-tv/35-mm-film-digital-Hollywood/*, accessed on 09/12/2012.

Asker, J. 2005. "Diagnosing Foreclosure due to Exclusive Dealing."

Berry, S., J. Levinsohn, and A. Pakes. 1995. "Automobile prices in market equilibrium." *Econometrica: Journal of the Econometric Society*, 841–890.

Berry, S., J. Levinsohn, and A. Pakes. 2004. "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market." *Journal of Political Economy*, 112(1): 68–105.

Berry, S.T. 1994. "Estimating discrete-choice models of product differentiation." *The RAND Journal of Economics*, 242–262.

Besanko, D., and M.K. Perry. 1993. "Equilibrium incentives for exclusive dealing in a differentiated products oligopoly." *The RAND Journal of Economics*, 646–667.

Brenkers, R., and F. Verboven. 2006. "Liberalizing a distribution system: the European car market." *Journal of the European Economic Association*, 4(1): 216–251.

Crawford, G.S., and A. Yurukoglu. 2012. "The welfare effects of bundling in multichannel television markets." *American Economic Review*, 102(2): 643–685.

Davis, P. 2006*a*. "Spatial competition in retail markets: movie theaters." *The RAND Journal of Economics*, 37(4): 964–982.

Davis, Peter. 2006*b*. "Measuring the Business Stealing, Cannibalization and Market Expansion Effects of Entry in the US Motion Picture Exhibition Market." *The journal of industrial economics*, 54(3): 293–321.

Dubin, Jeffrey A, and Daniel L McFadden. 1984. "An econometric analysis of residential electric appliance holdings and consumption." *Econometrica*, 345–362.

Einav, L. 2007. "Seasonality in the US motion picture industry." *The Rand journal of economics*, 38(1): 127–145.

Eliashberg, J., Q. Hegie, J. Ho, D. Huisman, S.J. Miller, S. Swami, C.B. Weinberg, and B. Wierenga. 2009. "Demand-driven scheduling of movies in a multiplex." *International Journal of Research in Marketing*, 26(2): 75–88.

Feige, U. 1998. "A threshold of ln n for approximating set cover." *Journal of the ACM (JACM)*, 45(4): 634–652.

Filson, D., D. Switzer, and P. Besocke. 2005. "At the movies: The economics of exhibition contracts." *Economic Inquiry*, 43(2): 354–369.

Gasparon, P, and B Visnar. 2005. "Coca-Cola: Europe-wide remedies in fizzy drinks."

Gil, R. 2009. "Revenue sharing distortions and vertical integration in the movie industry." *Journal of Law, Economics, and Organization*, 25(2): 579–610.

Gil, R., and W.R. Hartmann. 2007. "The role and determinants of concession sales in movie theaters: Evidence from the Spanish exhibition industry." *Review of Industrial Organization*, 30(4): 325–347.

Gil, Ricard, and Francine Lafontaine. 2012. "Using Revenue Sharing to Implement Flexible Prices: Evidence from Movie Exhibition Contracts." *The Journal of Industrial Economics*, 60(2): 187–219.

Grennan, M. 2013. "Price discrimination and bargaining: Empirical evidence from medical devices." *American Economic Review*, 103.

Hansen, L.P. 1982. "Large sample properties of generalized method of moments estimators." *Econometrica: Journal of the Econometric Society*, 1029–1054.

Hanssen, F.A. 2002. "Revenue-sharing in movie exhibition and the arrival of sound." *Economic Inquiry*, 40(3): 380–402.

Hicks, J.R. 1939. "Value and Capital: An inquiry into some fundamental principles of economic theory."

Hitsch, G.J. 2006. "An empirical model of optimal dynamic product launch and exit under demand uncertainty." *Marketing Science*, 25(1): 25.

Hochbaum, D.S. 1997. "Approximation algorithms for NP-hard problems." *ACM SIGACT News*, 28(2): 40–52.

Ho, J., K. Ho, and J.H. Mortimer. 2010. "Welfare effects of full-line forcing contracts in the video rental industry." *NBER Working Paper*.

Ho, K. 2009. "Insurer-provider networks in the medical care market." *The American Economic Review*, 99(1): 393–430.

Lee, R. 2012*a*. "Vertical Integration and Exclusivity in Platform and Two-Sided Markets." *New York University mimeo*.

Lee, Robin. 2012*b*. "Dynamic demand estimation in platform and two-sided markets." *New York University mimeo*.

Moretti, Enrico. 2011. "Social learning and peer effects in consumption: Evidence from movie sales." *The Review of Economic Studies*, 78(1): 356–393.

Mortimer, J.H. 2008. "Vertical Contracts in the Video Rental Industry." *Review of economic studies*, 75(1): 165–199.

Moul, C.C. 2007. "Measuring word of mouth's impact on theatrical movie admissions." *Journal of Economics & Management Strategy*, 16(4): 859–892.

Nevo, A. 2001. "Measuring Market Power in the Ready-to-Eat Cereal Industry." *Econometrica*, 69(2): 307–342.

OECD. 1996. "Competition Policy and Film Distribution."

Orbach, Barak Y, and Liran Einav. 2007. "Uniform prices for differentiated goods: The case of the movie-theater industry." *International Review of Law and Economics*, 27(2): 129–153.

Petrin, A. 2002. "Quantifying the Benefits of New Products: The Case of the Minivan." *Journal of Political Economy*, 110(4).

Scott-Morton, F. 2012. "Contracts that Reference Rivals." Presented at Georgetown University Law Center Antitrust Seminar.

Sunada, Mitsuru. 2012. "Competition among movie theaters: an empirical investigation of the Toho–Subaru antitrust case." *Journal of Cultural Economics*, 36(3): 179–206.

8 Appendix

8.1 Data

8.1.1 Attendance data

The attendance data set is adjusted to suit the discrete choice model used to explain consumer demand.

Pooling observations In the U.S. tickets to see movies in most movie theaters are "general admission", meaning they do not come with seat reservations. Abstracting from screen capacity constraints, it can thus be reasonably assumed that all tickets to see a given movie at a given time are viewed by consumers as identical, even if movies are played across a couple different screens. The discrete choice model is not well-suited to handling such multiple parallel screenings of one movie, since each movie/screening combo gets its own Ξ_{msct} draw, suggesting consumers get different levels of utility from seeing one movie on different screens. The discrete choice model will thus predict higher attendance if the movie theater puts on multiple screenings of the same movie at the same time, even if one screen could have handled all the demand.

In order to avoid this problem observations are pooled for a given movie in cases where screening times are close enough to be viewed by moviegoers as identical. The cutoff for time difference between screening which are pooled together is set at 60 minutes; this value was chosen to balance two goals. One one hand, a high value is needed to pool staggered releases of one movie, e.g. 6 screenings at 10-minute intervals between 12:00am and 12:50am. On the other hand, the value has to be low enough so as to differentiate between two sequential screenings of even the shortest movies on one screen e.g. a 7:10pm and 9:00pm screenings. Overall, the number of screenings is reduced 2.1%.

Eliminating observations This paper focuses on regular screenings of feature-length movies, which are the major source of revenues to exhibitors and distributors. In the data set, however, there are few observations that do not conform to this description. One way of dealing with them would be to leave them in the data set and estimate the coefficients driving their attractiveness to moviegoers. However, there are not enough observations for most of them to satisfy asymptotic requirements of the estimators. Thus the following screenings are removed from the data set:

1. Special screenings at non-standard times e.g. school trips in the morning hours

2. Free screenings e.g. summer movie series

3. Non-movie events e.g. NBA, NFL games, concerts

1.9% of observations are removed through this process .

8.1.2 Revenue split data

In order to simplify implementation and speed up model estimation and counterfactual simulation movies' revenue split values are "flattened" and discretized, while the range of possible values is limited for movies released on the break and in the second run.

Flattening Renting movies on a sliding scale is a practice that the industry has been moving away from and that is likely to be discontinued in the coming years, according to exhibitors. Overall, fewer than 19% of movies in the data set were rented on a sliding scale contract, a proportion which fell to below 15% in 2010 (the last full year of observation). For these reasons all movies in the counterfactual simulated are rented on a flat rate. In estimation, so as not to eliminate observations, the revenue split values for all movies rented on a sliding scale are "flattened" i.e. converted from a sliding scale to a flat rate. The new flat rate is calculated such that the total split of box office revenues between the distributor and exhibitor over the entire observed movie run is as close as possible to that under the original sliding scale pricing schedule.

Insert Figure 3 Here

Range limitation and value discretization As illustrated in Figure 3 the majority of revenue split values for movies on flat rate contracts released in the second run falls within the 35-55 range, with only one movie rented at 60. Additionally, the vast majority of movies are rented at revenue splits which are multiples of 5. Thus, to simplify and speed up calculations, revenue split values are allowed to only take of a limited number of discrete values from the set $[35, 40, 45, 50, 55]$. In estimation, revenue split values are modified to fit this set, with the new revenue split value being as close as possible to the observed value.

Insert Figure 4 Here

As illustrated in Figure 4 the majority of revenue split values for movies on flat rate contracts released on the break falls within the 50-60 range. Although for this set of movies the case is not as clear-cut as for movies released in the second run, here too revenue split values are discretized such that they fall into to the set $[50, 55, 60, 65]$.

8.2 Demand Model

8.2.1 Capacity constraints

The model does not explicitly take into account capacity constraints, however their impact on demand estimates is likely to be negligible, as in the sample less than 1% of screenings were sold out. One could expect, however, that consumers may choose to avoid going to screenings they *expect* could be close to sold out, even though they prefer this screening to all others, *ceteris paribus*. Not modeling this explicitly could lead to underestimating the value people place on the most popular screening periods (e.g. Friday evening) if there is a substantial number of screenings which are sold close to capacity. However, since in the sample less than 3% of screenings sell more than 75% of capacity any potential bias from this source is likely to be negligible

8.2.2 Initial moviegoing history

Consumer moviegoing history, ι_{it}, is a crucial determinant of demand as consumers in the demand model do not see the same movie twice. Since its value is not observed by the econometrician it needs to be simulated within the model, which creates the initial condition problem - what movies were seen by moviegoers prior to the first period of observation? In order to get around this problem the following approach is taken:

1. For each movie theater c the demand model is simulated for all periods $t \in T_c$ assuming $\iota_{i \min T_c} = \emptyset \; \forall \; i, c$, i.e. moviegoers had not seen any movies in periods prior to the first period of observation

2. Only time periods $t \geq \min T_c + t_{initial}$ $\forall c$ are taken into account when forming the GMM objective function, where $t_{initial}$ is set such that $\mathcal{M}^{+}_{cminT_c} \cap \mathcal{M}^{+}_{cminT_c+t_{initial}} = \emptyset$ i.e. none of the movies screened in the first period of observation were screened in period $minT_c + t_{initial}$, and any impact of the initial condition is second-order

Correspondingly, in counterfactual simulations only periods $minT_c + t_{initial}$ onwards are analyzed. Since $t_{initial}$ as defined above varies between simulations, its value is fixed at 5 such that for most simulations no movies screened in the first period of observation were screened in period $minT_c + t_{initial}$.

8.2.3 Controlling For Moviegoing History

The challenge of controlling for moviegoing history is similar to that faced by Lee (2012b), who has to account for consumers' ownership of gaming consoles at the same time as he models heterogeneous consumer tastes. His solution is to divide the consumer population into subgroups based on discretized values of the parameter capturing consumer heterogeneity, and keep track of the proportion of consumers who own gaming consoles within each group over time. This approach suffers from the curse of dimensionality, however, and limits the number of characteristics which can be used to explain heterogeneity in consumer tastes (Lee (2012*b*) only implements one dimension).

Instead, the approach taken in this paper relies on simulation techniques. For each movie theater c, in the first period $t = 1$ N individuals are drawn from the population (N = 500 in implementation), each with a full set of characteristics which may affect their moviegoing decisions. The algorithm then proceeds sequentially for each movie theater. In period t, given each individual i's moviegoing history ι_{it-1}, a BLP contraction mapping is used to back out δ_{ms}. Moviegoing decisions are then simulated for each individual i and recorded in $\{\iota_{it}\}_{\forall i}$, and the algorithm proceeds to period $t+1$.

8.3 Estimation and Identification

8.3.1 Conditionality of the primary set of moments

Unlike in the standard BLP setup, the expectation expressed in (15) is not unconditional - rather, it is conditional on the selection resulting from allowing consumers to see a given movie only once. In order to make sure the expectation holds, it is important to explicitly model this selection mechanism. The model does this by keeping track of movies seen by each one of the simulated individuals, $\{\iota_{mit}\}_{\forall m,i,t}$. When deciding which movies to see in period t, individual i only considers those for which $\iota_{mit} = 0$. At the end of each period the moviegoing of each individual i is simulated, setting $\iota_{mit'} = 1 \; \forall \; t' \geq t$ if he decides to see movie m in period t.

8.3.2 Identification

The estimation procedure exploits the panel nature of the dataset and employs numerous fixed effects:

1. *Screening period / movie theater / time interval* - fixed effects capture the observed utility from screening period $I_s^S(p)\beta^S$, the unobserved utility component μ_{sct}, as well as the price component $p_{sct}^A\beta^P$ (*time interval*

is defined such that over its duration ticket prices remain constant at a given movie theater c). The primary source of identification is the variation in sales as movies screened change between periods, but the screening periods remain constant for a given movie theater (variation in seasonal moviegoing demand is captured by the *week* fixed effects). If ticket prices change over the sample period the panel nature of the dataset allows for the identification of an additional set of fixed effects for this movie theater.

2. *Movie* - fixed effects capture the true movie quality, λ_m. Identification comes from two sources: time-variation in sales at a given movie theater as movies change, and from variation in sales across movie theaters whose choice of movies screened in a given time period differs.

3. *Week* - fixed effects capture the seasonal component of the unobserved utility, μ_w. Identification comes from time-variation in sales throughout the year.

4. *Year* - fixed effects capture the long-term trend component of the unobserved utility, μ_y. Identification comes from time-variation in sales over the years.

The coefficient on ticket price, β^P, is not separately identified. There are three places where it enters the estimation procedure, as per (16).The first, $p^A_{ic}\beta^P$, is captured along with the component of ΠD_i associated with the constant, which combined represent the additional utility children/seniors get from going to the movies *net* of the admissions price.[30] p^{3D}_c is captured along with β^{3D}_c by the dummy variable on whether a given screening is in 3D, x^{3D}_{msct} - this can be thought of as the utility consumers get from a given movie theater's 3D screening *net* of prices charged by this movie theater. Finally, as

[30] In implementation, due to the closeness in child/senior discounts across movie theaters and the added computational complexity of estimating four additional non-linear parameters, only one parameter is estimated to capture this effect

described above, $p^A_{sct}\beta^P$ is captured by the *screening period / movie theater / time interval* - fixed effects.

Identification for β^W comes from variation in the release date for a given movie across movie theaters in the sample e.g. if movie m was released on the break in one movie theater but in the second run in another movie theater. If the sample included a larger selection of movie theaters it should be possible to identify a β^W_m coefficient separately for each movie, however the limited number of movie theaters in the sample prevents this.

In the standard BLP setup observed heterogeneity in consumer tastes is identified using variation in consumer demographics between markets. Such variation is limited, however, for the markets in the sample. Instead, the primary source of identification are the micro-moments. Micro-moments #2, #3 and #5 help identify age-specific utility from going to the movies - the components of Π in (4) that correspond to the constant. Micro-moment #5, additionally, helps identify the remaining coefficients in Π: the utility derived by different age groups from movie characteristics such as genre or MPAA rating as well as from different screening periods.

Identification for the variance in consumer preferences comes from many sources. The first source of identification is the substitution patterns between products as these change across time periods. This helps identify the variance of ω_{im} but not the component of Σ that corresponds to screening periods, as these do not vary across time periods. The second source of identification is micro-moment #1 which captures the heterogeneity in moviegoing frequency over the course of a year and thus helps identify the component of Σ corresponding to the constant. Finally, as described in Lee (2012*a*) there is an additional source of identification that comes from the panel nature of the dataset and exploits the *self-selection* among consumers. For example, consider a world where the only potential source of heterogeneity between moviegoers is ω_{im}, and a situation where in week 2 of movie m_1's release the exhibitor releases another movie, m_2, with the same *mean* appeal for moviegoers accounting for attractiveness decay (i.e. $\lambda_{m_1} + \beta^W = \lambda_{m_2}$). If there is no heterogeneity in consumers' taste in movies ($Var(\omega_{im}) = 0$) then the decision which movie to see will be primarily driven by the idiosyncratic component ϵ_{imsct} — the model will then predict that, among the consumers who have not seen m_1 in week 1, as many will see m_1 as m_2 in week 2. If, instead, $Var(\omega_{im}) > 0$, than on average moviegoers who have not seen m_1 in week 1 will have a *lower-than-average* value of ω_{im_1} (i.e. $mean(\omega_{im_1}|\iota_{m_1 i1} = 0) < 0$), and thus in week 2 more of them will see m_2 than m_1.

8.3.3 Movie quality estimation: log-likelihood function

The log-likelihood function is the following:

$$\ell(\sigma^2_\mu, \{\sigma^2_{\nu c}\}_{\forall c}, \boldsymbol{\beta}^M, \{\hat{\lambda}_{0ct}, \lambda_{0t}\}_{\forall c, t \in T_c} | \{\lambda_m\}_{m \in \mathcal{M}^+}, x^M, \mathcal{M}^+, \mathcal{M}^-) = \sum_{\forall c} \sum_{t \in T_c} \ell_{tc}(\sigma^2_\mu, \sigma^2_{\nu c}, \boldsymbol{\beta}^M, \hat{\lambda}_{0ct}, \lambda_{0t} | \{\lambda_m\}_{m \in \mathcal{M}^+}, x^M, \mathcal{M}^+_{ct}, \mathcal{M}^-_{ct}) \tag{22}$$

where T_c is the set of time periods observed for exhibitor c, and $\lambda_{0t} = \max_{m \in \mathcal{M}^-(w_{mt}=0)}(\lambda_m)$ is the highest true movie quality value for movies not screened by any of the exhibitors.

The per-period log-likelihood function is:

$$\ell_{tc}(\cdot|\cdot) = \ln\Bigg(P_{m \in \mathcal{M}^-(w_{mt}=0)}\left(\hat{\lambda}_{0ct}, \lambda_{0t} | \boldsymbol{\beta}^M, \sigma^2_{\nu c}, \sigma^2_\mu\right) \prod_{m \in \mathcal{M}^+_{ct}(w_{mt}=0)} \phi_\mu(\lambda_m - \boldsymbol{x}^M_m \boldsymbol{\beta}^M) \Phi_\nu\left(\underline{\hat{\lambda}_{mc}}(\hat{\lambda}_{0ct}) \le \hat{\lambda}_{mc} \le \overline{\hat{\lambda}_{mc}}(\hat{\lambda}_{0ct})\right) \prod_{m \in (\mathcal{M}^+(w_{mt}=0) \setminus \mathcal{M}^+_{ct}(w_{mt}=0))} \phi_\mu(\lambda_m - \boldsymbol{x}^M_m \boldsymbol{\beta}^M) \Phi_\nu(\hat{\lambda}_{0ct} - \lambda_m) \Bigg) \tag{23}$$

where $P_{m \in \mathcal{M}^-(w_{mt}=0)}\left(\hat{\lambda}_{0ct}, \lambda_{0t} | \boldsymbol{\beta}^M, \sigma^2_{\nu c}, \sigma^2_\mu\right)$ is the joint probability that for $m \in \mathcal{M}^-(w_{mt} = 0)$: $\lambda_m \le \lambda_{0t} \forall m$, $\exists m$ s.t. $\lambda_m = \lambda_{0t}$, $\hat{\lambda}_{mc} \le \hat{\lambda}_{0ct} \forall m$, $\exists m$ s.t. $\hat{\lambda}_{mc} = \hat{\lambda}_{0ct}$ if $\hat{\lambda}_{mc} \ge \max_{m \in \mathcal{M}^-_{ct}(w_{mt}>0)} \left(\lambda_m + w_{mt}\beta^W_2\right)$. $\ell_{tc}(\cdot|\cdot)$ accounts for all movies released nationally in period t ($w_{mt} = 0$) and

1. never screened by any exhibitor (their λ_m is unknown to the econometrician): $\mathcal{M}^-(w_{mt} = 0)$

2. released this period (λ_m known, $\hat{\lambda}_{mc}$ within bounds calculated in stage 1): $\mathcal{M}^+_{ct}(w_{mt}=0)$

3. released in later periods (λ_m known, $\hat{\lambda}_{mc}$ below $\hat{\lambda}_{0ct}$) : $\mathcal{M}^+(w_{mt} = 0) \backslash \mathcal{M}^+_{ct}(w_{mt}=0)$

8.3.4 Movie quality estimation: algorithm

The algorithm

1. Consider one possible combination of $\sigma^2_\mu, \{\sigma^2_{\nu c}\}_{\forall c}, \boldsymbol{\beta}^M$ values

2. $\forall c, t$, determine $\boldsymbol{\lambda_0}$, the set of $\{\lambda_{0t}, \hat{\lambda}_{0ct}\}$ pairs which are consistent with observed schedules

3. calculate $\ell_{tc}(\cdot|\cdot)$ for all $\{\lambda_{0t}, \hat{\lambda}_{0ct}\} \in \boldsymbol{\lambda_0}$

4. set $t = \min_{\forall c}(T_c)$

5. find $\{\lambda_{0t}, \{\hat{\lambda}_{0ct}\}_{\forall c \text{ s.t. } t \in T_c}\}$ multiple which maximizes $\ell_{tc}(\cdot|\cdot)$

6. check that $\hat{\lambda}_{0ct}$ is consistent with movies released in previous periods i.e. $\hat{\lambda}_{0ct} \geq \max_{t'<t}\left(\lambda_{0t'} + (t-t') * \beta_2^W\right)\forall c$; if not, go to 7., else go to 8.

7. find $\left\{\hat{\lambda}_{0ct} \in \left[argmax\big(\ell_{tc}(\cdot|\cdot)\big), \max_{t'<t}\left(\lambda_{0t'} + (t-t') * \beta_2^W\right)\right]\right\}_{\forall c}$ which maximize $\sum_{\forall c}\sum_{t'=min_{\forall c}(T_c)}^{t} \ell_{t'c}(\cdot|\cdot)$

8. set $t = t+1$

9. iterate steps 4 - 7 while $t \leq max_{\forall c}(T_c)$, calculate $\ell(\sigma^2_\mu, \{\sigma^2_{\nu c}\}_{\forall c}, \boldsymbol{\beta}^M|\cdot)$ when done

10. iterate steps 1 - 8 until $\ell(\sigma^2_\mu, \{\sigma^2_{\nu c}\}_{\forall c}, \boldsymbol{\beta}^M|\cdot)$ is maximized

In point 2 inconsistent $\lambda_{0t}/\hat{\lambda}_{0ct}$ pairs are ones where:

1. $\hat{\lambda}_{0ct} > \min_{m \in \mathcal{M}^{+}_{ct}(w_{mt} \geq MRL_m)}$ i.e. the best alternative cannot be better than the worst of the movies being kept on whose minimum exhibition period is no longer binding (that is, the exhibitor is free to replace it but chooses not to do so), where MRL is the assumed minimum exhibition period restraint for movie m; this imposes restraints on $\hat{\lambda}_m$ values for movies released in period t and on λ_m values for movies released before t

2. $\hat{\lambda}_{0ct} < \min_{m \in \mathcal{M}^{+}_{ct}(w_{mt} \geq 1)} (\lambda_m + w_{mt}\beta_2^W)$ i.e. in period t when at least one of the movies that opens in movie theater c was released nationally before t the best alternative cannot be better than the worst movie that was actually released; this also implies...

3. ...$\lambda_{0t'} < \min_{m \in \mathcal{M}^{+}_{ct}(w_{mt} \geq 1)} (\lambda_m + w_{mt'}\beta_2^W)$ i.e. for a movie with known quality to be released it has to be better than similar alternatives in period t, and thus none of the movies released before period t can be better than this movie

These restrictions are exogenous to the estimation procedure and can be imposed by analyzing the exhibitor schedules.

The need for points 6-7 stems from the fact the, by definition,

$$\hat{\lambda}_{0ct} = \max \left[\max_{m \in \mathcal{M}^{-}_{ct}(w_{mt}=0)} \left(\hat{\lambda}_{mc}\right); \max_{m \in \mathcal{M}^{-}_{c}(w_{mt}>0)} \left(\lambda_m + w_{mt}\beta_2^W\right) \right]$$

can be driven either by the best movie released in period t that was not screened or the best movie released in previous periods that was not screened. By considering all $\{\lambda_{0t}, \hat{\lambda}_{0ct}\} \in \boldsymbol{\lambda_0}$ in point 3 the algorithm does not account for the latter component of $\hat{\lambda}_{0ct}$. If the resultant $\hat{\lambda}_{0ct} > \max_{t'<t} \left(\lambda_{0t'} + (t-t') * \beta_2^W\right)$ than it is consistent with the model. If the opposite is true, the algorithm needs to consider all candidate values between the two extremes and determine which maximizes $\ell(\cdot|\cdot)$ to this point.

Implementation details In the data set over 99% of movies are screened within the first 9 weeks of their nationwide release. In order to speed up calculations and better approximate actual exhibitor behavior the algorithm only allows movies to be considered for release up to 9 weeks after their nationwide release.

8.3.5 Counterfactual Calculations

Calculating $\boldsymbol{z}^*(\cdot)$ The problem of finding an optimal schedule when attendance at one screening depends on all other screenings can be viewed as a special case of the maximum coverage problem which is *NP-hard* i.e. no polynomial-time algorithm is known for solving it. The model employs the greedy algorithm, which has been shown to be the best polynomial-time algorithm to solve the maximum coverage problem (Hochbaum, 1997; Feige, 1998).

The aim of the algorithm is to allocate movies to empty screening-time/screen slots, the composition of which is identical to that actually observed, so as to maximize the movie theater's per-period profit function (9). This can be done with or without the no screen-sharing restriction. The algorithm puts together a movie schedule iteratively, at each step adding the movie/screening-time combination that most increases the *combined* profits, until there are no more empty slots to fill.

Capacity constraints need to be accounted for explicitly when calculating $z^*(\cdot)$. The algorithm assigns screens in decreasing order of capacity, thus ensuring that most attractive movie/screening period combinations chosen early on are assigned to the largest screens. For movie/screening period combinations where the predicted attendance exceeds the screen's capacity the algorithm considers assigning another screen to the movie - doing so does not expand the consumers' choice set, but instead raises the capacity constraint for this movie/screening period combination. After the algorithm is finished, attendance at each movie/screening period combination is capped at the combined capacity of assigned screens.

Calculating $\boldsymbol{r}^*_{mc}$ This algorithm aims to find $\boldsymbol{r}^*_{mc}$ which maximizes the distributor's profit from screening movie m in movie theater c: Π_{mc} (13). Relying on the fact that Π_{mc} is non-decreasing in $\boldsymbol{r}_{-mc}$ the algorithm starts with $\boldsymbol{r}_{mc} = min(R_{mt})\ \forall\ m \in \mathcal{M}^*_{ct}$, where R_{mt} is the range of values revenue splits can take.[31] It then proceeds to iteratively increase $\boldsymbol{r}_{mc}$ on each movie until no distributor finds it profitable to increase it any further.

[31]The set of possible revenue split values R_{mt} depends on whether the movie is released on the break or in the second run, and in both cases the values are limited to those actually observed in the sample

Table 1: Data set summary statistics, across movie theaters

	mean	minimum	maximum
Average annual attendance	84,197	33,015	174,131
Average annual # movies	78.5	63.6	86.4
Average # weeks on screens, by movie	2.9	2	3.8
Data period (months)	35.8	14	64
Screens per theater	4.3	3	6
Market size (local population)	13,134	4,380	44,737
Distance to closest competitor (miles)	48.4	23	116

Sources: data from movie theaters, Wikipedia, Google Maps

Table 2: Ticket prices by movie theater

Movie Theater	Base Price	Discounts / Premia			Base Price Changes	
		Matinee	Non-Adult	3D	Count	Magnitude
MT1	\$9.00	-\$2.00	-\$2.50	\$3.00	0	n/a
MT2	\$6.00	-\$1.00	-\$2.00	\$2.00	1	-\$0.50
MT3	\$7.75	-\$1.50	-\$2.25	n/a	0	n/a
MT4	\$7.75	-\$1.50	-\$2.25	n/a	3	\$0.75
MT5	\$8.00	-\$2.00	-\$2.50	n/a	1	\$0.50

Notes: prices shown at the end of the sample period (Jan 2011); discounts and premia reported relative to base price; "non-adult" refers to child and senior discounts; base price changes reported over the sample period

Source: exhibitors

Table 3: Average proportion of new movies to all movies screened

	all weeks	restricted: weeks ahead of event		
		1	2	1 - 2
Blockbuster movies	52.4%	55.0%	52.2%	53.2%
High-attendance weeks	52.4%	51.6%	54.0%	52.9%
Holidays	52.4%	58.7%	55.3%	56.3%

Notes: Each row represents an event exhibitors might be expected to take into account ahead of time when scheduling movies; "all weeks" column represents average for whole data set, while "restricted" represents average ahead of each event

Table 4: Demand model coefficients

			Interaction with age dummy variables			
Variable	Means	St Dev	2-11	12-24	25-59	60+
Constant	-10.01***	1.00***		1.00***		
	(2.14)	(0.13)		(0.17)		
3D	0.26*					
	(0.19)					
On the break ($w_{mt} = 0$)	0.42***					
	(0.09)					
w_{mt}	-0.03***					
	(0.01)					
MPAA: PG-13/R			-0.75			
			(3.16)			
Genre: Family/Animated			1.00			
			(2.01)			
Genre: Fantasy/SciFi/Animated						-5.75**
						(1.75)
Screening: Weekdays before 5pm				-3.75*	-3.75*	
				(2.30)	(2.30)	
ω_{im}		0.89***				
		(0.13)				
Fixed effects	movie-, week-, year-, screening time/movie theater/time interval-					
Number of observations	54,785					

Note: ***, **, and * indicate significance at the 1 percent, 5 percent and 10 percent levels, respectively

Table 5: Top/Bottom 10 Movie Fixed Effects

Top 10	Fixed Effect	BOR ($MM)	Bottom 10	Fixed Effect	BOR ($MM)
Pirates of the Caribbean 4	3.75	423.3	Ninja Assassin	-1.65	38.1
Cars	3.13	244.1	Saw 3D: the Final Chapter	-1.67	45.7
Wild Hogs	3.04	168.3	Death At a Funeral	-1.70	42.7
Night At the Museum	2.91	250.9	She's Out of My League	-1.71	31.6
The Chronicles of Narnia	2.82	291.7	Machete	-1.82	26.6
Over the Hedge	2.70	155.0	Skyline	-1.87	21.4
X-men: the Last Stand	2.70	234.4	The Next Three Days	-1.89	21.1
Superman Returns	2.66	200.1	The Crazies	-1.98	39.1
Shrek the Third	2.60	320.7	Why Did I Get Married Too?	-2.13	60.1
Pirates of the Caribbean 2	2.58	309.4	Case 39	-2.16	13.2

Note: BOR is the total U.S. Box Office Revenue over the course of a movie's theatrical run

Table 6: β^M estimates

Variable		Estimate	Standard Error	Significance
β^M	const	-2.00	(0.38)	***
	IMDB rating	0.12	(0.05)	***
	MPAA G-rating	0.80	(0.20)	***
	Budget	5.00	(0.56)	***
	Genre: fantasy	0.40	(0.19)	**
	Genre: sci-fi	0.40	(0.17)	***
	Genre: musical	-1.00	(0.31)	***
	Genre: documentary	-0.80	(0.31)	***
σ_μ^2		2.00	(0.39)	***
$\sigma_{\nu 1}^2$		0.25	(0.14)	***
$\sigma_{\nu 2}^2$		0.50	(0.25)	***
ℓ		1,331.02		

Notes: [(a)] Budget is expressed in \$100M; ***, **, * indicate significance at the 1 percent, 5 percent and 10 percent levels (standard error are calculated using bootstrapping)

Table 7: Results of Counterfactual Simulations

MT1

Scenario:	1. Base	2. Base	3. No screen-sharing		4. Screen-sharing allowed	
	Absolute	Absolute	Absolute	Change	Absolute	Change
Attendance (thousands)	231,872	242,777	252,019	3.8% to 8.7%	254,595	4.9% to 9.8%
# movies screened	107	127	146	15.0% to 36.4%	153	20.5% to 43.0%
Consumer utility (utils)	1,412	1,467	1,546	5.4% to 9.5%	1,576	7.4% to 11.6%
Exhibitor profits (thousand $)	1,189	1,277	1,349	5.6% to 13.4%	1,359	6.4% to 14.3%
Distributor profits (thousand $)	1,008	1,023	1,041	1.8% to 3.2%	1,056	3.2% to 4.7%
Total Welfare Change				3.9% to 8.7%		5.0% to 9.9%

MT2

Scenario:	1. Base	2. Base	3. No screen-sharing		4. Screen-sharing allowed	
	Absolute	Absolute	Absolute	Change	Absolute	Change
Attendance (thousands)	57,618	60,868	64,985	6.8% to 12.8%	66,773	9.7% to 15.9%
# movies screened	94	132	188	42.4% to 100.0%	214	62.1% to 127.7%
Consumer utility (utils)	4,506	4,816	5,006	3.9% to 11.1%	5,275	9.5% to 17.1%
Exhibitor profits (thousand $)	300	317	345	9.0% to 15.1%	361	14.0% to 20.3%
Distributor profits (thousand $)	259	273	283	3.9% to 9.5%	285	4.5% to 10.1%
Total Welfare Change				6.5% to 11.9%		9.4% to 14.9%

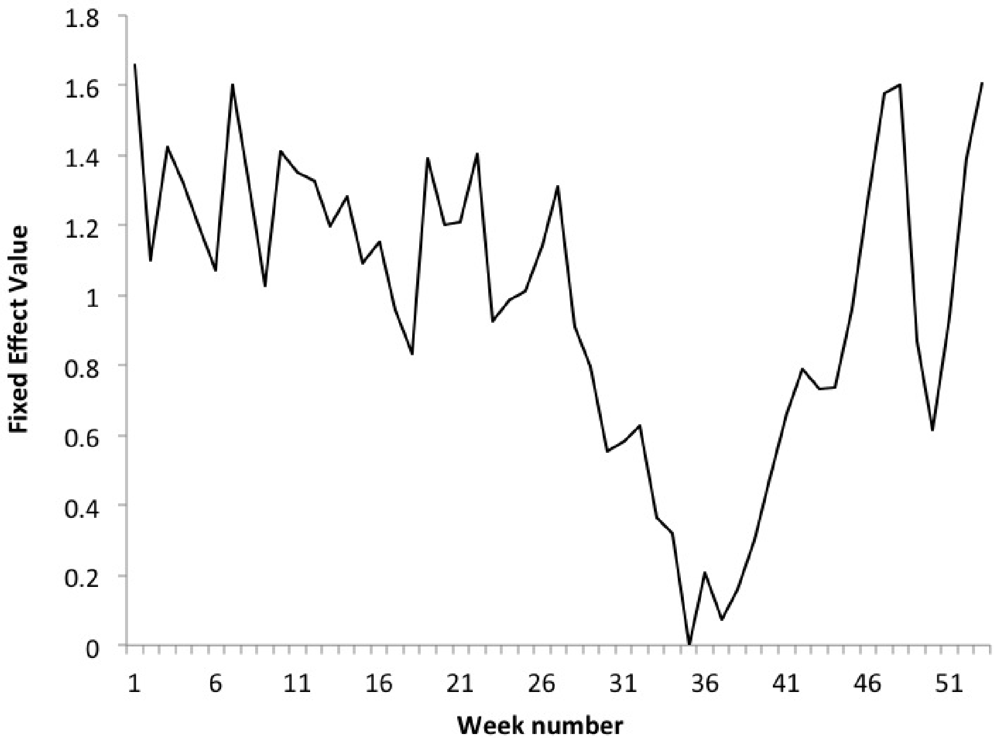
Notes: MT1 has six screens while MT2 has three screens. Abs represents absolute values, while Δ represents change relative to base cases. In scenarios 3. and 4. percentage change is expressed relative to scenarios 1. and 2. , which provide a lower and upper bound on the expected gains from removing the restraints. See section 5.2 for detailed explanations of each experiment. Total Welfare Change is the sum of change in consumer utility expressed in $ terms, as described in Section 4.1, as well as changes in exhibitor and distributor profits.

Table 8: Deviation payoffs at potential equilibria

Strategy Payoffs	1. All impose restraints gains from lifting restraints		2. No one imposes restraints gains from imposing restraints	
	MT1	MT2	MT1	MT2
Major #1	-21.2% to -13.9%	-12.4% to -8.5%	17.9% to 26.2%	5.9% to 10.9%
Major #2	-16.7% to -14.1%	-32.4% to -21.1%	6.0% to 11.4%	11.2% to 29.6%
Major #3	-25.8% to -8.0%	-31.1% to -15.8%	-1.4% to 6.8%	11.6% to 30.9%
Major #4	-16.2% to -7.6%	-16.4% to 2.5%	14.2% to 24.1%	14.3% to 36.4%
Major #5	-40.5% to -15.2%	-48.1% to -33.1%	14.3% to 31.4%	47.6% to 93.0%
Major #6	-34.0% to -12.0%	-20.1% to -19.3%	24.4% to 47.6%	31.7% to 43.2%
Average	-25.7% to -11.8%	-26.7% to -15.9%	12.6% to 24.6%	20.4% to 40.7%

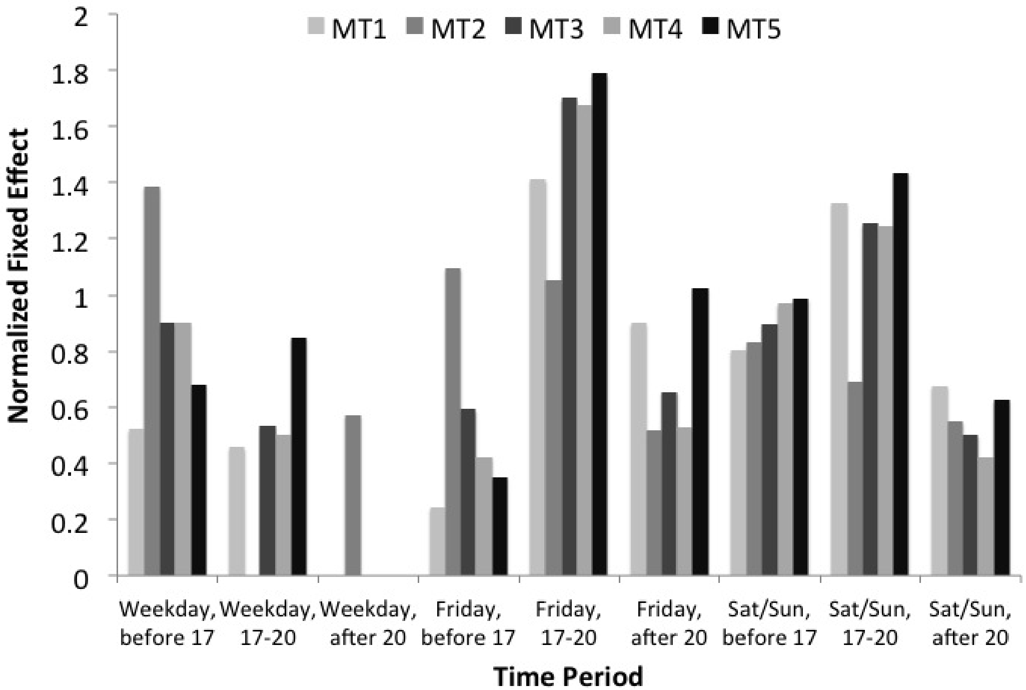
Notes: See Section 5.3 for a detailed explanation of the setup. Each row represents one of the six Major distributors, with the entries representing the distributor's change in profits from lifting or imposing restraints. Two scenarios are considered: 1. where all distributors impose restraints and 2. where no distributors impose restraints; the benefits from deviating from the strategy are modeled for MT1 and MT2 separately.

Figure 1: Week fixed effects



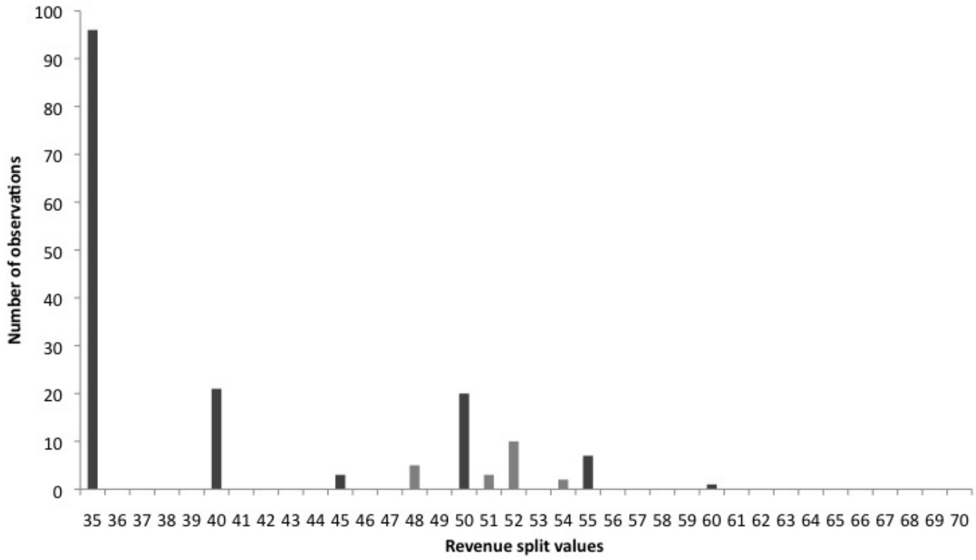
Note: values are normalized such that the smallest value equals zero

Figure 2: Movie theater fixed effects



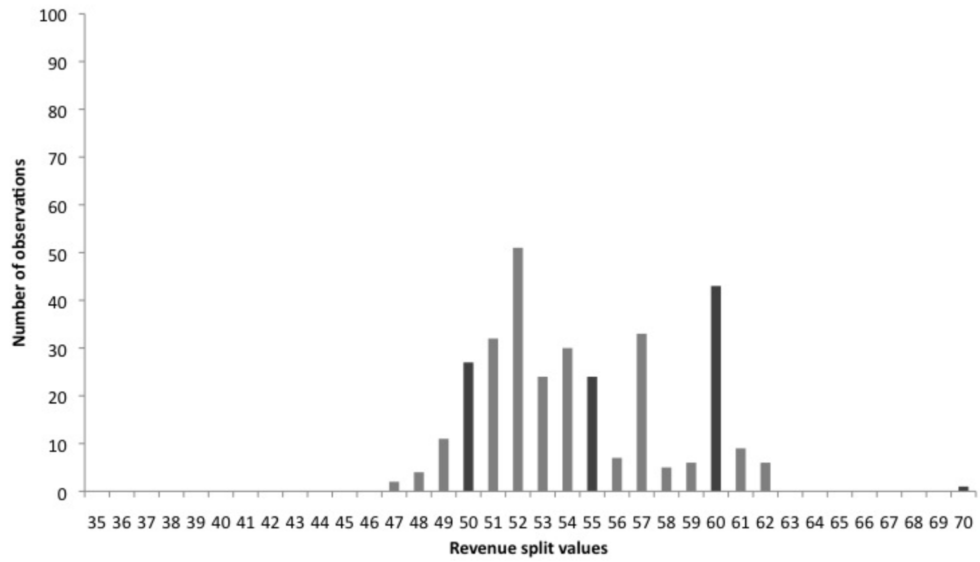
Note: values are normalized such that the smallest value equals zero

Figure 3: Revenue split values, movies on flat rate contracts released in the second run



Source: revenue split data from movie theaters

Figure 4: Revenue split values, movies on flat rate contracts released on the break



Source: revenue split data from movie theaters