NOT ALL RIVALS LOOK ALIKE: ESTIMATING AN EQUILIBRIUM MODEL OF THE RELEASE DATE TIMING GAME

LIRAN EINAV*

I develop a new empirical model for discrete games and apply it to study the release date timing game played by distributors of movies. The results suggest that release dates of movies are too clustered around big holiday weekends and that box office revenues would increase if distributors shifted some holiday releases by one or two weeks. The proposed game structure could be applied more broadly to situations where competition is on dimensions other than price. It relies on sequential moves with asymmetric information, making the model particularly attractive for studying (common) situations where player asymmetries are important. (*JEL* C13, C51, L13, L15, L82)

> . . .a very serious game of strategy is at work–a cross between chess and chicken–which studio distribution chiefs play year round, but with increasing intensity during the summer and holiday release period. (*New York Times,* December 6, 1999)

> Hubris. Hubris. If you only think about your own business, you think, "I've got a good story department, I've got a good marketing department, we're going to go out and do this." And you don't think that everybody else is thinking the same way. In a given weekend in a year you'll have five movies open, and there's certainly not enough people to go around. (Joe Roth, chairman of Walt Disney Studios, answering a question about the large number of major movies opening within days of each other; *Los Angeles Times,* December 31, 1996)

I. INTRODUCTION

The number of Americans who go to the movies varies dramatically over the course of the year, and sometimes more than doubles within a period of two weeks. At the same time, the first week accounts for almost 40% of the box office revenues of the average movie. The combination of these two facts makes the timing of launching a new movie a major focus of attention for distributors of movies. With virtually no subsequent price competition, the movie's release date is one of the main short-run vehicles by which studios compete with each other.

In this paper, I develop and estimate a model of discrete games, which allows me to analyze this release date timing game. Most empirical industry studies focus on price and quantity competition, taking other product characteristics as given; in many industries, however, prices play a very small role, and competition is channeled through other product attributes. The entertainment industry is a prime example; competition among movies, television programs, or Broadway shows is on nonmonetary product attributes, such as content, advertising, and time. Therefore, understanding competition in such industries requires a model that endogenizes some of these product attributes, especially those that can be changed in the short run. This paper provides a framework in which release decisions

*This paper partially originates from chapter 1 of my 2002 Harvard University dissertation, and I am particularly grateful to Ariel Pakes for his guidance and support in the early stages of this paper. I thank three anonymous referees for many useful suggestions, and Susan Athey, Pat Bajari, Lanier Benkard, Tim Bresnahan, Peter Davis, Mike Mazzeo, Aviv Nevo, Peter Reiss, Katja Seim, and seminar participants at the Society of Economic Dynamics 2003 annual meeting in Paris, the Econometric Society 2004 Winter meeting in San Diego, UW Madison, and the University of Tokyo for useful comments on earlier drafts. I thank Oren Rigbi for superb research assistance, Jeff Blake (*Sony Pictures*), Ron Kastner (*Goldheart Pictures*), and Terry Moriarty (*Hoytz Cinemas*) for insightful conversations, and *ACNielsen EDI* and *Exhibitor Relations Inc.* for helping me obtain important portions of the data. I acknowledge financial support from the National Science Foundation and the Stanford Institute for Economic Policy Research.

Einav: Associate Professor of Economics, Stanford University, 579 Serra Mall, Stanford, CA 94305-6072; and Research Associate, National Bureau of Economic Research, 1050 Massachusetts Avenue, Cambridge, MA 02138. Phone 650-723-3704. Fax 650-725-5702, E-mail leinav@stanford.edu.

Economic Inquiry
(ISSN 0095-2583)
Vol. 48, No. 2, April 2010, 369–390

doi:10.1111/j.1465-7295.2009.00239.x
Online Early publication September 28, 2009

can be endogenized and nonprice competition can be analyzed.

The absence of price competition is useful because it allows the focus of the analysis to be on the timing dimension without relying on assumptions about the nature of the postrelease price competition.[1] Together with the frequent timing decisions associated with different movies, this makes the motion picture industry quite attractive for empirical analysis of the timing game. This industry, however, is not the only example in which timing decisions play a central role. Similar timing considerations are also important in the release decisions of books, compact disks, and other new products, as well as in the scheduling decisions of major events, television programs, flight schedules, and promotional sales.[2]

When taking on a new project, distributors of movies typically plan for a Friday release, which falls within a relatively short window of time carefully chosen to match the type of the movie. This makes the choice of the release date a very discrete one. The exact Friday within the season is generally determined by a strategic timing game played among distributors. Before each release season, distributors scramble to release their movies on big holiday weekends, when demand is high. When doing so, each distributor tries to release on the attractive holiday weekend and at the same time to deter competitors from doing the same. The extent to which distributors are successful in doing so largely depends on the quality of the movies at their disposal, and on the way they can compete with movies released by other studios. Thus, modeling this timing game must account for the asymmetries among movies, and for the variation in these asymmetries across release seasons.

Specifically, I develop and estimate a sequential-move game with private information; I assume that the observed release date decisions are the equilibrium outcome of such a game. The empirical analysis relies on data from the U.S. motion picture industry between 1985 and 1999. I take the season in which a movie is released as given, and focus the analysis on the strategic decision of the specific release date within the season. To specify payoffs, and in particular to assess the heterogeneous substitution effect in demand between movies, I rely on the estimates of the demand for movies obtained in a companion paper (Einav 2007). In that paper, I modeled demand for movies as a function of movie quality, decay in the demand for a movie, and seasonal underlying demand for movies. The main focus of the analysis in the current study is on evaluating the extent to which distributors of movies over- or underestimate this underlying demand vis-à-vis the substitution effect from competing movies. I find that the release pattern implies that underlying demand for movies is much more seasonal than is estimated by the demand system. That is, the results suggest that the release dates of movies are too clustered on holiday weekends and that distributors could increase theatrical revenues by shifting holiday release dates by one or two weeks. Different possible explanations, such as uncertainty and conservatism, are discussed.

Beyond this specific application, the game structure I develop is an important contribution of this paper because it is likely applicable more generally. The model builds on ideas from the existing literature on discrete games, but combines these ideas together in a new way, thereby providing certain attractive properties.[3] The model is a sequential-move game with asymmetric information. Under certain restrictions on the private information, one can construct a relatively simple pseudo-backward induction algorithm to solve for the unique perfect Bayesian equilibrium of the game. For any payoff structure, the sequential structure of the game leads to a unique probability distribution over all possible outcomes, thus allowing for a simple maximum likelihood estimation. The private information assumption facilitates evaluation of the likelihood function by avoiding the difficulties (due to complex regions of integration) that would arise with complete information.

An attractive property of the proposed model is its ability to accommodate an unrestricted payoff structure. In particular, the model and its

1. The fact that ticket prices hardly vary across seasons and movies is taken as given throughout this paper. This is an interesting puzzle, which is discussed in more detail in Orbach and Einav (2007).

2. There are only a few papers that analyze competition in time. They mostly use reduced-form statistics to assess the equilibrium outcome (Borenstein and Netz 1999; Chisholm 1999; Corts 2001; Simonsohn 2008). Goettler and Shachar (2001) construct a strategic scheduling game between television networks, but do not use it for estimation. Sweeting (2008), who models the timing of radio advertising, is a notable exception.

3. See Section II for more details. See also Reiss (1996), Einav and Nevo (2006), and Draganska et al. (2008) for related reviews and discussions of existing models of discrete games.

estimation could accommodate player identities and asymmetries in substitution patterns. Symmetry assumptions, which are crucial for many existing empirical models of discrete games, are not necessary. Such symmetry restrictions often seem implausible; in a wide range of industries, decision makers care not about the number of competitors they face but also about competitor identities. All else equal, a software developer is more likely to enter a market in which another small software company operates rather than a market in which Microsoft participates. Similarly, a movie distributor would rather release his movie during the same week as a low-budget movie release than during the same week as the Star Wars release. In accordance with the above observations, studying differentiated product markets and the varying degrees of substitution among products has proven important in demand estimation.[4] The model I propose allows for the incorporation of players' identities into models of discrete games and could facilitate research that combines entry and location games with empirical demand models; thus far, these two strands of the literature have evolved quite independently of each other.

The rest of the paper is organized as follows. Section II presents the model, its properties, and its estimation; it also describes its relationship to the existing literature of discrete games and illustrates the differences. Section III describes the industry and the data, and Section IV presents the empirical specification and the results. Section V concludes.

II. THE MODEL

A. The empirical model

Let the set of players in market m be $i = 1, 2, \ldots, N_m$ and the discrete (finite) action space for player i be A_i^m. Given the actions of all players (for simplicity of notation, the m subscripts are suppressed), $a \in A_1 \times A_2 \times \cdots \times A_N$, payoffs for player i are given by

$$\pi_i(a; X, \beta, \eta) = \hat{\pi}_i(a; X, \beta) + \epsilon_{a_i}^i, \tag{1}$$

where β is a vector of parameters and $\varepsilon_{a_i}^i$ is an i.i.d (across actions and players) draw from a type I extreme value distribution with a precision parameter η.[5] The vector of $\varepsilon_{a_i}^i$'s is assumed to be *private information* of player i. The private information can be thought of as nonstrategic considerations that may make a certain player more likely to choose a certain action, regardless of the actions of the other players. It could also be thought of as optimization errors. The magnitude of the estimated η provides an indication for the explanatory power of the model. This is because the variance of the unexplained portion of the payoffs, $\varepsilon_{a_i}^i$, is decreasing in η, thus η provides a measure of the explanatory power of the deterministic part of the payoffs.[6] The higher η is the more we can explain the observed decisions by the estimated payoffs rather than by the random error. An insignificant estimate of η implies that the model for the payoffs has no statistically significant explanatory power.[7]

This specification leads to simple logit probabilities. Conditional on the other players' decisions, a_{-i}, movie i chooses action a_i with the following probability:

$$\Pr(a_i|a_{-i}) = (\exp(\eta\hat{\pi}_i(a; X, \beta))) \Bigg/ \left(\sum_{a_i' \in A_i} \exp(\eta\hat{\pi}_i(a_i', a_{-i}; X, \beta))\right). \tag{2}$$

The game is played sequentially with each player moving exactly once according to a prespecified order, which is known to the players but may be unknown to the econometrician. The solution concept is a perfect Bayesian equilibrium. Note that the payoffs of each player i depend only on the action taken by the other players, but not on the realizations of their $\varepsilon_{a_j}^j$'s ($j \neq i$). Therefore, each player's strategy depends only on the actions of players who moved previously, but not on their exact draws

4. See, among many others, Berry, Levinsohn, and Pakes (1995) and Nevo (2001).

5. This distribution has a cdf $F(x) = \exp(-\exp(-\eta x))$, mean γ/η (where $\gamma = 0.577$ is the Euler's constant), and variance $\pi^2/6\eta^2$. As η goes to 0, the variance goes to infinity, and as η goes to infinity, the variance goes to 0.

6. To empirically identify η, one must pin down the level of $\hat{\pi}$ through any other assumption. For example, if $\hat{\pi} = X\beta$, it is easy to see that η cannot be separately identified from β. In this paper, η is identified because external data is used to estimate $\hat{\pi}$ and pin down its level. More generally, doing so requires some external information that would pin down the level of one of the other parameters of the model.

7. McKelvey and Palfrey (1998) provide the more general properties of such games, and use it to analyze experimental data. This literature uses the term *quantal response equilibria* to describe such games.

of ε's. This assumption greatly simplifies equilibrium analysis because it implies that from the player's perspective, opponent types are irrelevant when opponent actions are known. Consequently, given the prespecified order of play, the game can be solved backwards in a simple way.

In what follows, I outline the simple algorithm that is used to solve the model. In the rest of the paper, I refer to this algorithm as *pseudo-backward induction*. Given N players, let the order of play be given by a permutation $o \in \mathcal{P}_N$, such that $o(m) = j$ implies that player j is the mth player to move. Let $\text{prev}(j) = \{k : o^{-1}(k) < o^{-1}(j)\}$ denote the set of players who play before player j. I solve the game backwards. The last player to move, $o(N)$, conditions on the other players' decisions, $a_{-o(N)}$. Therefore, we can use Equation (2) to see that $a_{o(N)}$ is chosen with probability

$$\Pr(a_{o(N)}|a_{-o(N)}) = (\exp(\eta\hat{\pi}_{o(N)}(a_{o(N)}, a_{-o(N)}; \beta))) \Big/ \left(\sum_{a'_{o(N)} \in A_{o(N)}} \exp(\eta\hat{\pi}_{o(N)}(a'_{o(N)}, a_{-o(N)}; \beta))\right). \tag{3}$$

Going backwards, we can update the continuation values for all other players by integrating out over player $o(N)$'s decision probabilities, namely,

$$\hat{\pi}_i^{N-1}(a_{-o(N)}; \beta) = \sum_{a_{o(N)} \in A_{o(N)}} \Pr(a_{o(N)}|a_{-o(N)}) \hat{\pi}_i(a_{o(N)}, a_{-o(N)}; \beta) \quad \forall i \in \text{prev}(o(N)), \tag{4}$$

and

$$\pi_i^{(N-1)}(a_{-o(N)}; \beta, \eta) = \hat{\pi}_i^{(N-1)}(a_{-o(N)}; \beta) + \varepsilon^i_{a_i} \quad \forall i \in \text{prev}(o(N)). \tag{5}$$

The key is that the $\varepsilon^i_{a_i}$'s are invariant to a_{-i}, they depend only on a_i, so they can be taken out of the sum. These modified payoffs would directly enter the decision of player $o(N-1)$. They are also updated for the rest of the players because we use these modified payoffs below as part of the iterative procedure. In particular, we can use Equation (2) again, but with respect to the modified payoffs, which are given by Equation (4). This procedure can be done iteratively up to the player who moves first, with each iteration being the following:

$$\Pr(a_j|a_{\text{prev}(j)}) = (\exp(\eta\hat{\pi}_j^{o^{-1}(j)}(a_j|a_{\text{prev}(j)}))) \Big/ \left(\sum_{a'_j \in A_j} \exp(\eta\hat{\pi}_j^{o^{-1}(j)}(a'_j|a_{\text{p}(j)}))\right), \tag{6}$$

and

$$\hat{\pi}_i^{(o^{-1}(j)-1)}(a_{\text{prev}(j)}; \beta) = \sum_{a_j \in A_s} \Pr(a_j|a_{\text{prev}(j)}) \hat{\pi}_i^{(o^{-1}(j))}(a_j, a_{\text{prev}(j)}; \beta) \quad \forall i \in \text{prev}(j). \tag{7}$$

Thus, in the end of the procedure we obtain a probabilistic equilibrium play for each player, and hence a strictly positive probability measure over each potential outcome of the game. In particular, given an order o, the likelihood of observing an outcome a is given by

$$\Pr(a|o) = \prod_{j=1}^{N} \Pr(a_j|a_{\text{prev}(j)}, o). \tag{8}$$

Given a probability measure over all possible permutations (orders of play), the unconditional likelihood of observing an outcome a is given by

$$f(a) = \sum_{o \in \mathcal{P}_N} \Pr(o)\Pr(a|o) = \sum_{o \in \mathcal{P}_N} \left[\Pr(o) \prod_{j=1}^{N} \Pr(a_j|a_{\text{prev}(j)}, o)\right]. \tag{9}$$

If a natural order of play exists and is observed, one can use this order and think about o as being observable. Alternatively, one can assume a uniform distribution over all permutations, i.e., $\Pr(o) = 1/N!\forall o \in \mathcal{P}_N$. Finally, when the data are rich enough or when there are enough restrictions on the payoff structure, one can estimate the probabilities of different permutations. Although this is not implemented in the application below, a simple yet general way to specify the probability measure over the order permutations is to assign a commitment measure for each player, given by $\mu_j = W_j\delta + \zeta_j$, where ζ_j is distributed i.i.d extreme value, W_j is a vector of observed characteristics of player j which affect his commitment power, and δ is a vector of parameters that can be estimated. The order of moves is then dictated by the commitment measure μ_j. This implies that the probability of an order o is given by

$$\Pr(o) = \prod_{j=1}^{N} (\exp(W_j\delta)) \Big/ \sum_{k \notin \text{prev}(o(j))} \exp(W_k\delta). \tag{10}$$

Given M distinct and independent markets and a specification for $\hat{\pi}_i(a; X, \beta)$, the model can be estimated using maximum likelihood.

Finally, it is important to note what would be altered in the model if we considered a perfect information game, that is, a game of the same structure in which all $\varepsilon^j_{a_j}$'s are common knowledge. The key difference, for the econometrician, is that the observed decision of, say, the player who moves last provides information about that player's $\varepsilon^j_{a_j}$'s. In the perfect information case, unlike in the derivation above, the other players know these $\varepsilon^j_{a_j}$'s and hence such information must be taken into account by the econometrician when assigning probabilities to the other players' decisions at earlier points of the game. Thus, these decisions would have to be analyzed in light of a truncated extreme value distribution, for which we do not have closed-form solutions. To address this in a useful way, we will need to employ simulation estimators, which will solve for the subgame perfect equilibrium for each set of simulated vectors of $\varepsilon^j_{a_j}$'s.[8] This complication is the reason why the imperfect information case is computationally more attractive.

B. Relationship to the Literature

Existence and uniqueness are typically the two properties of equilibrium we analyze. Empirically, we generally assume that the data are generated by an equilibrium behavior, thus eliminating any existence problem. Multiplicity, however, remains a major issue. In order to understand the estimation problems associated with multiplicity of equilibria, let us denote the empirical model by $y(X_i, \varepsilon, \beta) \subset Y$, that is, a mapping from the model primitives, namely, observables X, unobservables ε, and parameters β, into the model predicted outcome $y(X_i, \varepsilon, \beta)$, which is a subset of all potential outcomes Y.

If $y(X_i, \varepsilon, \beta)$ is a singleton for all X_i's, β's, and ε's (i.e., equilibrium is always unique) then estimation is straightforward: one can proceed with, say, maximum likelihood estimation, with the likelihood of the observed outcome, y_i, given by $\Pr[\varepsilon \in \{\varepsilon | y_i = y(X, \varepsilon, \beta)\} | X, \beta]$. If, however, $y(X_i, \varepsilon, \beta)$ is nonunique then such a likelihood-based estimation procedure cannot be carried out unless we extend the model to have an additional assumption about the probability measure over the set of possible outcomes, $y(X_i, \varepsilon, \beta)$. Several comments are in place: (1) in general, theory tells us nothing about these equilibrium selection probabilities; (2) to be specified correctly, one has to account for all possible equilibria, for any given $(X_i, \varepsilon, \beta)$; (3) sometimes we may tell a story why one equilibrium is more likely than another, and this could be thought of as an extension of the model, which essentially provides uniqueness.

An approach taken by several authors (Berry 1992; Bresnahan and Reiss 1990) and more recently generalized by Davis (2006b) is to set up a model that does not provide uniqueness, but provides the econometrician with a coarser partition of the empirical model, which satisfies uniqueness. For example, in the context of entry models, these authors show that while, given $(X_i, \varepsilon, \beta)$, the model may have multiplicity of equilibria, all such equilibria share a common feature, which is the number of entering firms. Thus, the econometrician can condition on the number of entering firms, but not on their identities, and then apply likelihood-based (or other) estimation techniques. This is a coarser partition of Y in the sense that different observations are treated the same. While this approach has proven useful, it has two main limitations. First, the approach is not efficient in the sense that it treats different observables in the same way and hence does not use all the information provided by the data.[9] Second, a more important limitation is that strong symmetry assumptions must be imposed to get a unique prediction in models with more than two players. For example, in entry models, firms' payoffs are assumed to be invariant to permutations of the entry decisions made by their opponents. These assumptions seem quite unrealistic for a wide set of applications in which entrants are not drawn at random but are endogenously

8. The appendix of Berry (1992) conceptually describes such a simulation estimator. With player asymmetries, however, the procedure described there would be computationally more intensive.

9. Indeed, Tamer (2003) proposes a more efficient estimator, which exploits the additional information provided by the data.

drawn from a well-defined population of heterogeneous firms. Therefore, for many research questions, these models may prove unsatisfactory and may alter the economic implications of the results. Mazzeo (2002) relaxes this symmetry assumption by introducing different types of products, and conditioning the analysis on the number of entering firms of each type. The main restriction still remains: all potential entrants are ex-ante identical, and profits of a player are invariant to a permutation of his opponents' type choices. Moreover, extending Mazzeo's model to more than two or three types is computationally infeasible. Thus, the main two limitations remain largely unaddressed.

Two recent alternative approaches to deal with multiple equilibria have been developed. First, several papers (Andrews, Berry, and Jia 2007; Beresteanu, Molchanov, and Molinari 2008; Ciliberto and Tamer 2008) show that in the presence of multiplicity of equilibria, one can place bounds on the parameters of interest rather than obtain point estimates for them. The potential of these methods has yet to be fully realized, especially when there exists important variation in observable characteristics across observations. A second approach (Aguirregabiria and Mira 2007; Bajari, Benkard, and Levin 2007; Pakes, Ostrovsky, and Berry 2007) uses a two-step estimation procedure to get around the multiplicity problem. This approach assumes that the game has a reduced form, thereby avoiding the multiplicity problem. It relies heavily on accurate (nonparametric, ideally) estimation of the policy functions, which are then used to back out the structural parameters. While useful in many settings, this approach requires either large data sets or a small set of state variables. Many of the typical data sets and applications in industrial organization (for which the current paper is an example) do not satisfy either of these requirements.

Finally, a somewhat more structural approach is to change the structure of the game in such a way that equilibrium would be unique.[10] Bresnahan and Reiss (1990) and Berry (1992) suggest ways to do this by imposing a sequential structure on the game, which yields a generically unique subgame perfect equilibrium. This full information version, however, becomes computationally unattractive as we relax symmetry assumptions and increase the dimensionality of the game. Seim (2006) enriches the game structure by moving to games with asymmetric information. This makes the strategy of each player simpler from the econometrician's point of view because it now depends only on the firm-specific unobserved variables rather than on the whole set of unobservables in the market. Indeed, Seim (2006) is able to find a unique Bayesian Nash equilibrium and use it for estimation. Several limitations remain. First and foremost, the equilibria in such games are not necessarily unique.[11] Second, the search for the equilibrium strategies must involve an intensive numerical search for a fixed point, thus making computational complexity increase quite rapidly with the dimension of the problem. Third, just as in Mazzeo (2002), the same symmetry assumptions discussed above are still present: all opponents are ex-ante identical.

The model developed in the previous section is therefore in the spirit of Seim (2006), but with a sequential structure, as in certain specifications of Bresnahan and Reiss (1990) and Berry (1992). The (standard) assumptions on the private information structure guarantee uniqueness of equilibrium and imply that the equilibrium can be found using a pseudo-backward induction algorithm, thus alleviating some of the computational burden present in other models. Therefore, it incorporates different existing ideas into a game structure that guarantees uniqueness, is not restricted by symmetry assumptions, and is computationally attractive. In entry games, for example, such structure should be particularly attractive for situations in which additional information on postentry values is available (e.g., Berry and Waldfogel 1999, Orhun 2005, Ellickson and Misra 2007, or Watson, 2009). Such information would typically make symmetry assumptions internally inconsistent, and it would provide valuable insight on the structure of postentry values, which can be easily incorporated into the game structure just outlined.

10. Within this class, I also consider imposing a predefined probability distribution over the different equilibria, as in, for example, Bajari, Hong, and Ryan (2008). We can just think of an additional latent variable (the outcome of the "public randomization device"), conditional on which equilibrium is unique.

11. Seim (2006) numerically shows that there is a unique symmetric equilibrium for her particular model and data. More generally, however, there are no assumptions about the model that can guarantee uniqueness. Moreover, once rivals are allowed to be asymmetric, we would not be able to focus on symmetric equilibria, thereby the scope of finding multiplicity of equilibria would be even greater.

C. A Simple Illustration

Here I use a simple two-player entry game to illustrate the way the model works and how its predictions compare with those of other models used in the literature. The key (conceptual and computational) advantages for using this model only show up when we extend the game to include more actions and a greater number of (potentially asymmetric) players. Thus, this illustration is aimed to provide intuition about the mechanism and prediction of the model but not to highlight the computational advantages.

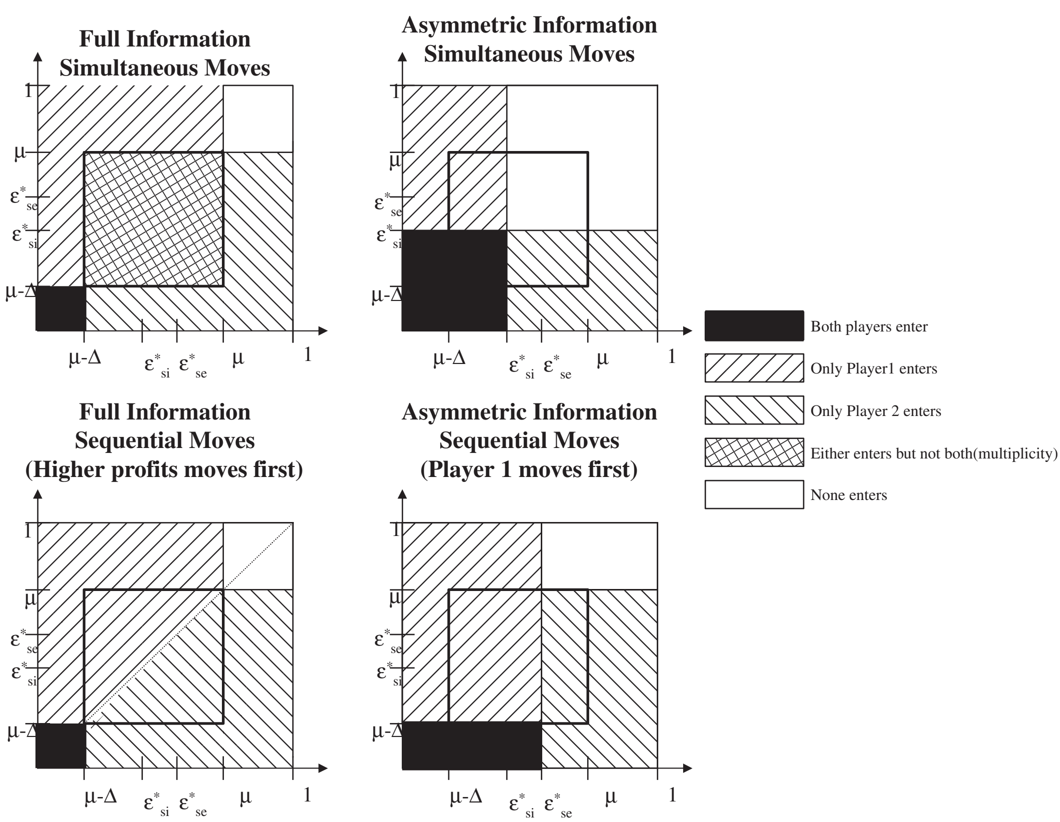
The example consists of a simple entry game. Each of the two players has to decide whether to enter the market or not. If player $i = 1, 2$ stays out, he obtains payoffs of 0. If he enters, he pays (sunk) entry costs of ε_i and collects payoffs of μ if his opponent stays out, and $\mu - \Delta$ otherwise (with $\mu > \Delta > 0$). For simplicity, assume also that ε_1 and ε_2 are both drawn independently from a uniform distribution over [0, 1].

We will consider different assumptions on the order of play and on the information structure. In all of these cases, we assume that the ε's are unknown to the econometrician, and that μ and $Delta$ are the estimable parameters of interest. Thus, I am interested in comparing how the different assumptions give rise to different probability distributions over outcomes. There are two types of information structures: a full information game, in which both ε's are known to the players, and an asymmetric information case in which each player only knows his own ε. Players either move simultaneously or sequentially.

Figure 1 shows the different cases. The upper left panel shows the full information simultaneous game, which is analyzed in Bresnahan and

FIGURE 1

Equilibrium Prediction under Various Modeling Assumptions



Notes: This figure shows the predicted outcomes of different game structures in a simple symmetric two-player entry game. For further discussion, see Section IIC.

Reiss (1990) and in Berry (1992). The square in the middle is the area that gives rise to multiplicity of equilibria. As pointed out in these papers, once one allows sequential structure, equilibrium is unique. For any point at the "multiplicity region," the player who moves first is the only player who enters in equilibrium. The sequence can be determined outside the model, so that the whole "multiplicity region" is allocated to one outcome, or, as suggested in Berry (1992) and in Mazzeo (2002), one can assume that the higher profit player moves first. The latter case is shown in the bottom left panel of Figure 1: within the multiplicity region only player 1 enters to the left of the 45° line, and only player 2 enters to the right of the line.

Consider now the case in which entry cost is private information. In the simultaneous-move case, as in Seim (2006), equilibrium follows a cutoff point strategy for each player; if ε_i^* is player i's cutoff point, his strategy is to enter if and only if $\varepsilon_i < \varepsilon_i^*$. The Bayesian Nash equilibrium in this simple example is given by the solution to the following two equations: $\varepsilon_1^* = \mu - F(\varepsilon_2^*)\Delta$ and $\varepsilon_2^* = \mu - F(\varepsilon_1^*)\Delta$ where $F(\cdot)$ is the cdf of ε. Once we impose the uniform distribution we obtain a unique equilibrium, in which the symmetric cutoff point is $\varepsilon_{\text{sim}}^* = \mu/(1+\Delta)$. The distribution over outcomes is depicted in the upper right panel of Figure 1.

Finally, the case of sequential moves with asymmetric information, which is the model used in this paper, is shown at the bottom right panel of the figure. It shows the distribution of outcomes when player 1 is the first mover (the case for player 2 being the leader is symmetric). Under the assumptions, the second mover just follows his full information strategy, conditional on the action played by player 1 (the first mover). Player 1 foresees this and uses a cutoff point strategy for entry, which is the solution to $\varepsilon_1^* = \mu - F(\mu - \Delta)\Delta$. With uniform distribution we obtain $\varepsilon_{\text{seq}}^* = \mu - (\mu - \Delta)\Delta$. It is easy to see that $\varepsilon_{\text{seq}}^* > \varepsilon_{\text{sim}}^*$; knowing that his action will be observed by player 2, player 1 can use it to be more aggressive in equilibrium.

Several comments are in place. First, in the full information case, moving first is advantageous (at least in this simple two-player entry game). In contrast, once information is asymmetric, there are cases in which moving first is a disadvantage. Consider, for example, a case in which ε_1 and ε_2 are just below μ. In such cases, the second mover will be the one entering the market and making positive profits. This is because the asymmetric information creates a trade-off: the first mover has a commitment power, but he also faces uncertainty. The second mover, in contrast, has no information problem: once his opponent has already moved, knowing his opponent's entry cost has no additional value. Second, as a consequence of the sequential moves, the likelihood of ex-post regret is much lower when compared to the simultaneous move case. Ex-post regret is experienced whenever a player would have liked to reverse his own action, once his opponent's action has been revealed. In Figure 1, regret is experienced in all areas in which the black and white rectangles on the right differ from those on the left. It is easy to see that, in most cases, these areas are much smaller in the sequential-move case. This is just a direct consequence of the previous argument: with sequential moves, only the first mover can experience regret, while the second mover effectively has no information problem. This also illustrates why I view the sequential game with asymmetric information as somewhat in between the two versions—with complete and incomplete information—of simultaneous move games. In particular, this is true once we randomize over the identity of the player who moves first.

D. Remarks

Regret. Empirical models with asymmetric information are vulnerable to the regret critique. The argument is that the asymmetric information may give rise to outcomes which would not be sustainable in the long run, as the players would like to change their previous actions. In the entry game illustrated above, for example, this happens when ε_2 is sufficiently high and ε_1 is just below μ. In both versions of asymmetric information, none of the players enter in equilibrium. Once player 1 finds out, however, that player 2 does not enter, player 1 would have liked to reverse his action and enter the market. The sequential move structure partially addresses this critique. As mentioned, the players who move late are less prone to information problems and hence less likely to experience regret. Thus, in general, the likelihood of regret is smaller under the sequential structure.

More importantly, the regret critique is more relevant for entry games than for other location choice games. If one interprets a choice as sinking a location-specific cost, then the regret argument has no bite. While, ex-post, a player would have liked to change his action, he has already sunk his choice-specific cost, so reversing it is costly. The entry story is a somewhat unique example in which a regret critique is more valid: it is more difficult (although possible) to think of irreversibilities associated with the choice of staying out of a market. For other sets of potential actions, irreversibility is much more plausible. In particular, this is the case in the application used in this paper; if choosing a particular release date for a movie implies sinking date-specific costs (e.g., printing posters or buying television advertising slots just before the release date), then the regret critique is less relevant.

Computation. Given the parameters of the model, there are two separate computational burdens. The first is to compute the entries in the payoff matrix, namely, to compute the postentry payoffs for each player, for any potential equilibrium outcome. If each of the N players has K actions to choose from, one needs to compute NK^N numbers (and repeat it for any value of the parameters). This may be computationally intensive if the parametric form of payoffs is both fully flexible and has a nontrivial functional form. Such computational issues do not arise in the existing literature, where symmetry assumptions imply much smaller sets of different entries. In the extreme symmetric case, where firms are identical, all we need is to compute N different numbers. Thus, it is important to emphasize that such a computational limitation, which arises from relaxing the symmetry assumption (and will be binding in the present application), is unrelated to the specific game structure that is being estimated.

Given the payoff matrix, the second computational burden is to compute the distribution over potential equilibrium outcomes implied by the model. To address this issue, the empirical model proposed here may be quite useful compared to others proposed in the literature (e.g., Seim 2006). The pseudo-backward induction algorithm is computationally linear in the number of players for any given order of play. Thus, one need not rely on numerical search routines, the computation time of which is typically hard to bound. There are, of course, $N!$ different orders of play to check, but this still gives the econometrician a clear bound on the computation time. In addition, if solving the model for all different orders of play is the computational bottleneck in a given application, it is quite easy to set up a simulated likelihood estimator, which will simulate a smaller number of order permutations, and will solve the model only for this smaller subset of games. Finally, as described below, one can impose other restrictions on the order of play that may be computationally more attractive.

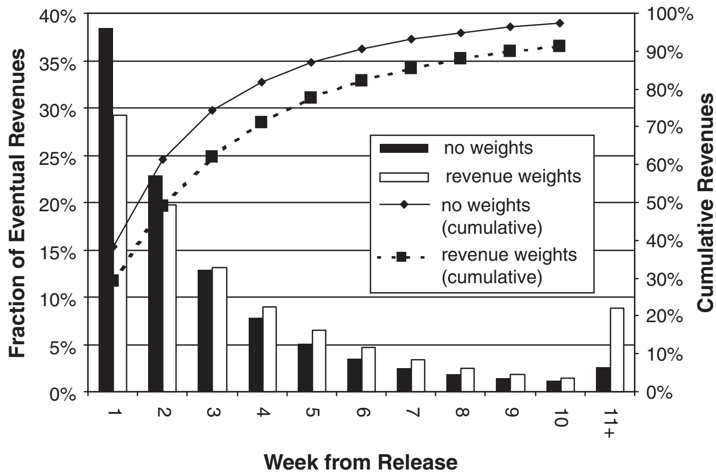
The Order of Moves. Clearly, once the model is of sequential structure, the order of moves is important. As already mentioned, however, it is somewhat less important once asymmetric information is present: in such a case moving first is not always an advantage. Moreover, I conjecture that in a large set of applications the qualitative results regarding the economic parameters of interest would not be very sensitive to the specific assumptions about the order of play. This is, at least, what I find in the current application.[12]

There are several different types of assumptions one could make about the order of play. First, by imposing more symmetry assumptions across players one needs to check less permutations because different orders of play would give rise to the same distribution of outcomes. For example, in the case of symmetric firms, there is only one order to check for. Second, one can either assume a uniform random order across different permutations of players, so each order is chosen with probability $1/N!$, or alternatively use a parametric family of distributions over permutations, one of which was proposed in the end of Section IIA. I do not attempt the latter in the present application; the identification of such parameters is more likely to be possible if we either put more structure on payoffs, or if we find variables which affect "commitment power" but do not enter the payoff function.[13] Finally, in many applications, one can use external information and impose it on the order of play. For example, the historical order of entry as in Toivanen and Waterson (2005), or the sequence of initial release date announcements in the current context, often allows the data to provide a natural order. Ordering moves by the size or quality of the players is also a reasonable assumption (Quint and Einav 2005). In general, once players are asymmetric, we gain

12. This is also related to Mazzeo (2002), who finds that different assumptions on the game structure had a very small effect on his result.

13. This would be an exclusion restriction. For a similar argument in a similar setup, see Bajari, Hong, and Ryan (2008).

FIGURE 2
Distribution of Movie's Box Office Revenues over Its Life Cycle



Notes: The "no weights" series calculates weekly percentages for each film separately, and then applies simple averages of these percentages over all movies. The "revenue weights" series calculates a weighted average, where the weights are proportional to the total box office revenues of each movie. This figure shows the distribution of total box office revenues over the movie's life cycle. The bars stand for the week-by-week share, while the lines stand for the cumulative share as of the end of the corresponding week. It can be seen that most of the revenues are concentrated in the first few weeks, with the first week accounting, on average, for almost 40% of the eventual box office revenues, and the first four weeks accounting for about 80% of them. Once I weight the averages by the gross box office revenues of the different films (white bars and dashed line), the distribution is less skewed and has a wider tail, suggesting that revenues of bigger movies decay slower.

more player-specific information and hence can use this information to determine the order of play in a more natural way.

III. INDUSTRY AND DATA

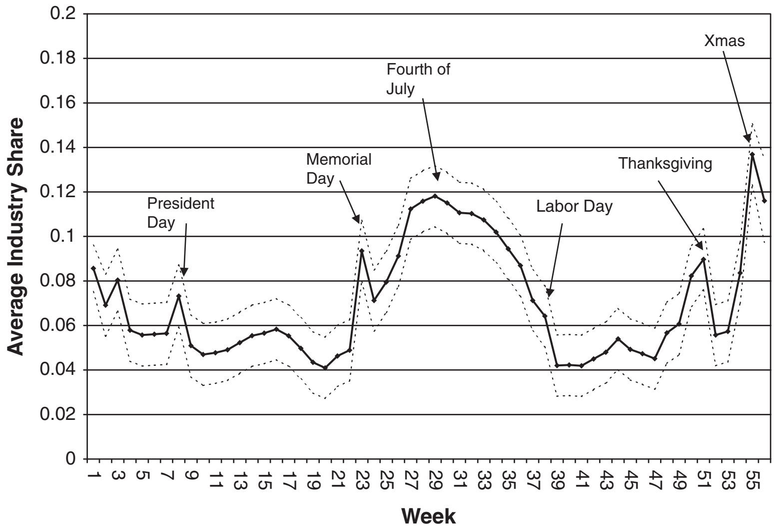
The distributors of motion pictures are those in charge of taking the movie from the end of the production stage to the theaters. This is done typically by the distribution arm of the major studios, as described in more detail in Einav (2002) and in the references therein. One of the main strategic decisions made by distributors is the movie release date. The two important considerations factored into this decision are the strong seasonal effects in the demand for movies and the competition that will be encountered throughout the movie's run. Typically, movies with higher expected revenues are released on higher demand weekends, so there is a trade-off between the seasonal and the competition effects. The importance of the release date is greatly magnified by the fact that the performance during the first week accounts for a sizeable amount of the overall performance of the movie. On average, box office revenues in the first week account for almost 40% of the total domestic revenues (Figure 2).[14] An additional reason for the importance of the release date choice is the view that high revenues in the first week create information and network effects which increase revenues in subsequent weeks.[15]

Figure 3 presents the strong seasonality in the industry, plotting weekly average industry revenues (normalized by ticket prices and the size of the U.S. population). Major holidays such as Memorial Day, Fourth of July, Thanksgiving, Christmas, and New Year's are historically associated with strong box office performance. Consistent with this revenue pattern, the conventional wisdom is that box office revenues are strong throughout the summer season and during the Christmas winter holiday period. The period following Labor Day up to mid-November is

14. Furthermore, about 70% of the weekly revenues are collected in the weekend.

15. To quote from Lukk (1997): "In this business, if you are not the number one film the week you are open, you usually are never the number one." See also Moretti (2007) and Moul (2007).

FIGURE 3
Seasonal Effects in Total Admissions



Notes: This figure shows the seasonality in total industry sales. The vertical axis ("industry share") is the industry's weekly revenues, normalized by average ticket price and by the U.S. population. Thus, it can be thought of as the per capita number of movies seen each week (an industry share of, say, 0.1 implies that 1 of every 10 people in the United States goes to the movies in the corresponding week). The figure shows the industry shares, averaged over the 1985–1999 sample period. The figure clearly demonstrates the perceived seasonal effects in the industry. The year has two strong periods, the long summer period (Memorial Day to Labor Day) and the Christmas Winter holiday period. The Spring and the Fall are typically considered very weak periods for the industry, and the drop after Thanksgiving is generally seen as a "shopping period." The dashed lines stand for deviations of two standard errors. The number of weekly dummies is 56 to account for the timing variation in U.S. holidays across the years (see Einav 2007 for details).

considered to be very weak, as is the period from the beginning of March to mid-May.

The identity of the competing movies is the second consideration taken into account when setting the release date. Distributors are wary of releasing a movie in close proximity to another movie with which competition will be strong. Furthermore, even once release dates are set, distributors often change them in response to new information concerning release dates of similar movies chosen by other distributors (see in more detail later on). Another strategy practiced by distributors is to announce their movie's release data early with the hope that preemptive action will deter other distributors from choosing the announced date. This practice is especially common with movies that are widely expected to be successful.

I use two distinct data sets for this paper. The first contains detailed information about all movies domestically released between 1985 and 1999 and is described in detail in the companion paper (Einav 2007). There I use these data to obtain demand estimates. Some of the estimates from the demand system are used in the current paper as an input into the empirical model proposed above. To motivate the applicability of the empirical model described above to the release date decision, I collected a second data set. This is a unique data set regarding the prerelease information about scheduled release dates, describing the dynamic process that leads to the eventual schedule. The source of the data is the "Feature Release Schedule," which is published monthly by *Exhibitor Relations Inc*.

In the beginning of each month, the publication lists the updated release schedule of all movies that are in the making but have not been released as of yet. Typically, movies are first listed about 12–18 months before their scheduled release. At this stage, many of the movies are in the process of casting or are in early stages of production. Thus, when first entering the monthly report, movies are generally not assigned to a specific release date. Rather, they are given a more general release season, such as "Summer 2002," "Christmas 2002," or just as "coming." As the scheduled release approaches, the release date becomes more specific, for example, "Late Summer 2002" or "Early July 2002," converging eventually to a specific date.[16]

The data cover roughly all the titles that were eventually released between 1985 and 1999, a total of 3,363 titles. To get an idea of what the data look like, let me use Bruce Willis's *Die Hard: With a Vengeance* (aka *Die Hard 3*) as an example. It was first listed as "May 1995" in the September 1994 issue of the publication. In December 1994, the schedule became more specific—May 12, 1995—but a month later it was pushed back by two weeks, to May 26, 1995 (Memorial Day). In February 1995, the movie's release was moved again, to May 19, 1995, which was the eventual release date. The sequence of announcements for the 1999 release of *Star Wars: The Phantom Menace* was less eventful; it was first listed as "May 1999" in the issue of May 1998. In the September 1998 issue, the announcement became more specific, May 21, 1999, and remained the same until its actual release.[17]

A major characteristic of the data is the frequent changes in the release schedule of certain movies. This is somewhat surprising, given the costs associated with changing a release date. Such costs are incurred for several reasons, such as committed advertising slots, the implicit costs of reoptimizing the advertising campaign, reputational costs, etc. The costs become higher as the changes in release date are done closer to the scheduled release. While some of these changes are the result of unforeseen production delays,[18] most of these changes are made for strategic reasons, and may provide some indication of unobserved characteristics of the movie, such as quality and commitment power. Supporting this idea, industry practitioners and the popular media describe the scheduling game as a war of attrition.

16. Agency issues provide an additional incentive for early announcements of release dates. The director typically edits the film until the very last day before the release, so the announced release date is used to set a final deadline to the production process.

17. In fact, in the April 1999 issue, the May 21 (Friday) announcement was changed to May 19 (Wednesday). However, as will be discussed later, I tabulate dates at the weekly level, making these two dates effectively identical.

18. See Einav and Ravid (2007) for analysis of such schedule changes.

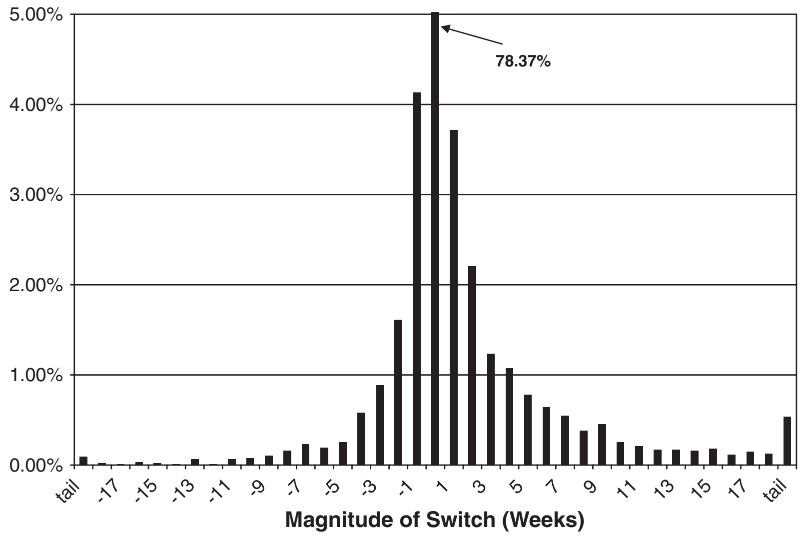
Across all movies and announcements, more than 20% of the monthly announcements are changes in relation to the most recent announcement of the same film. Moreover, more than 60% of the movies changed their release dates at least once. Figure 4 provides the distribution of the magnitude (in weeks) of these changes. The distribution of these changes is roughly symmetric, and the majority of changes shift the release date by a small number of weeks; 75% of the observed changes do not shift the release date by more than a month. Both the symmetry of the distribution and its shape indicate that it is unlikely that the majority of changes are made for exogenous nonstrategic reasons, such as production delays. The likelihood of a movie changing its release date is not significantly correlated with the movie's size, measured by its production cost. However, movies with higher box office quality (as estimated in Einav 2007) are significantly less likely to change release dates. One interpretation of this is that movies' estimated quality is originally highly correlated with their production cost, but as the shape of the finished product becomes clearer, films that turn out to be potential disappointments shift away from their previously announced release dates.

This pattern of frequent small switches seems consistent with the idea that movies, in general, are produced with a target season in mind, while the "fine-tuned" choice of the precise release date within the season is subject to more strategic consideration. Because over 75% of the movies are released on Fridays, and an additional 20% on Wednesdays, it seems natural to think of the release decision as a discrete choice among a small number of alternatives. Such a pattern lends itself nicely to the empirical model described in the previous section: at the end of the production stage each movie is scheduled to release during a specific season, while the exact release week is the outcome of a strategic timing game played against all other movies released during the same season.

While the true timing game is probably best approximated by a repeated announcement game with increasing switching costs (see Caruana and Einav 2008 for a formal analysis of such games), using such games for estimation is computationally infeasible. Therefore, the once-and-for-all sequential-move game, as proposed

FIGURE 4

Distribution of the Magnitude of Switches in Announced Release Dates



Notes: This figure is based on 1,897 movies, covering movies that were eventually released nationwide between 1985 and 1999. The figure shows the distribution of the magnitude of switches of announced release dates. A switch is a change in the announced release date compared to the most recent announcement of the same movie (which is generally made a month before). The distribution is taken over all movies and announcements, and tabulates the difference (in weeks) between the new announcement and the previous one. A difference of 0 implies no change. A positive difference implies a shift forward of the release date, and a negative difference implies making the release date earlier than announced before. This figure provides two main insights. First, the distribution is roughly symmetric (with a somewhat fatter tail in the positive part, for obvious reasons). Second, the majority of the changes shift the release date by a small number of weeks. These two observations suggest that these changes are done mainly for strategic reasons, and not because of exogenous factors, such as production delays. Note that the bar at 0 is out of scale, and accounts for 78% of the announcements.

in Section II, may be viewed as a reasonable alternative.[19]

IV. SPECIFICATION AND RESULTS

A. Overview

In the companion paper (Einav 2007), I estimate demand for motion pictures, where the weekly demand for a movie is driven by three components: the quality of the movie, the decay in quality since the movie's release, and the underlying seasonal pattern. Using a simple nested logit specification (with one nest for all movies, and a second nest for the outside good), the weekly market share of movie j during week t is given by

$$s_{jt} = (D_t^{\sigma} + D_t)^{-1} \exp((\theta_j - \lambda(t - r_j) + \tau_t + \xi_{jt})/(1 - \sigma)), \quad (11)$$

where θ_j is the movie quality,[20] r_j is the release date (in weeks) of movie j, τ_t is the underlying level of demand in week t, ξ_{jt} is a disturbance term which reflects the deviation from the common decay pattern, D_t is given by

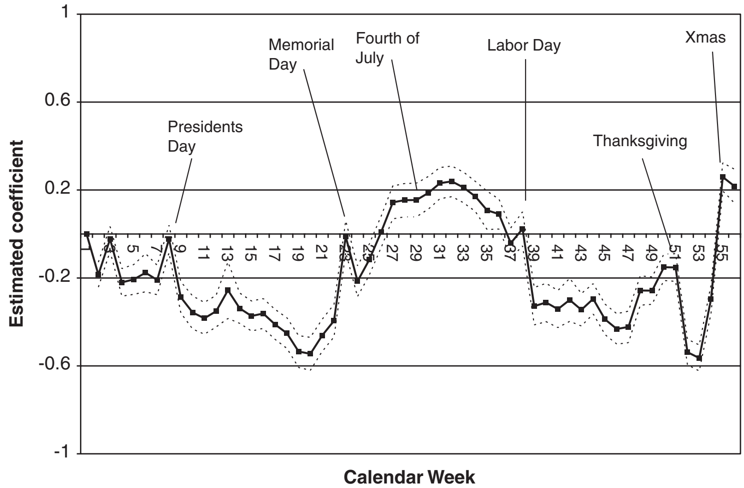
$$D_t = \sum_{k \in J_t} \exp((\theta_k - \lambda(t - r_k) + \tau_t + \xi_{kt})/(1 - \sigma)), \quad (12)$$

and J_t is the set of movies in theaters during week t. An important finding of Einav (2007) is that the estimates for underlying seasonality

19. One can think about the once-and-for-all assumption as if switching costs are insignificant early on, but become very high at a certain point in time. The order of moves assumed for the sequential game is just the ordering of the points in time at which these jumps in the switching costs occur.

20. One should think of quality as reflecting attractiveness or "box office appeal," which is not necessarily related to cinematographic quality.

FIGURE 5
Seasonal Effects in Underlying Demand



Notes: This figure plots the estimated coefficients on the weekly dummy variables from estimating the nested logit movie demand model of Einav (2007). There are two major differences compared to the seasonal pattern of industry revenues (Figure 3). First, the seasonal variation is smaller, suggesting that about one-third of the seasonal variation is explained by variation in quality. Second, the seasonal pattern is slightly different, with, for example, the end of the summer looking relatively much better than industry revenues indicate. The dashed lines stand for deviations of two standard errors. The estimated decay coefficient λ is -0.22 (with a standard error of 0.014) and the estimated substitution coefficient σ is 0.524 (0.030). The number of observations is 16,103 (1,956 movies). The number of weekly dummies is 56 to account for the timing variation in U.S. holidays across the years (see Einav 2007 for details).

are somewhat different from the conventional wisdom, as reflected by Figure 3. I reproduce these estimates for underlying seasonality in Figure 5.

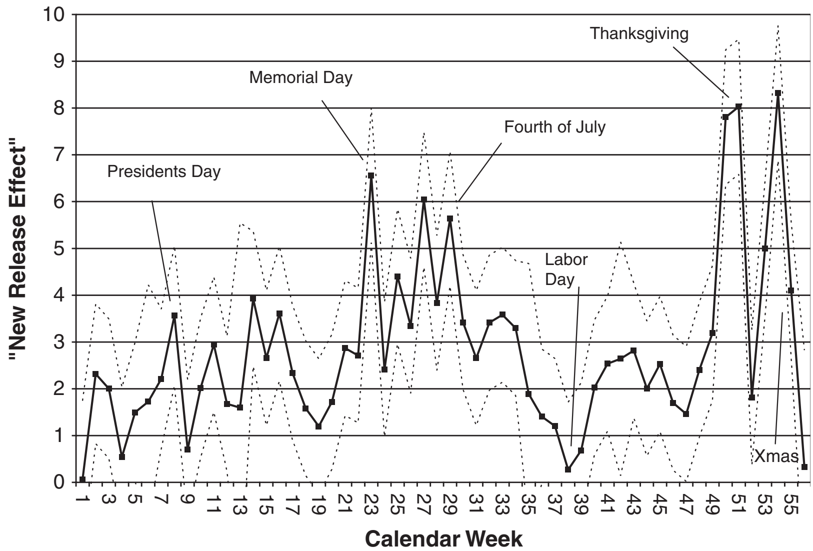
Simple analysis may suggest that, taking the estimates for underlying seasonality as given, distributors do not make their release decisions according to these estimates. The seasonal release pattern is described in Figure 6, showing that many of the top movies are released on a few big holiday weekends. The current application complements these findings by addressing two key issues. First, it examines the within-season variation in the release pattern, addressing a concern that the choice of a season may be driven by other omitted factors.[21] Second, it accounts for strategic effects by using the empirical model developed in this paper.

21. For example, many movies may release early in the summer in an attempt to leave enough time to make the high-demand video and DVD season around Christmas. Certain movies may also cater to a specific target audience, and characteristics of moviegoers change across seasons. All these factors are less relevant when considering the choice of a specific week within a given season.

The estimation strategy is to take the demand estimates of movie quality and decay pattern as given, and to estimate the underlying demand parameters from the game, that is, from the observed release pattern. The focus on the underlying demand is for several reasons. First, the other parameters of the demand system are less controversial, and hence make it a less interesting exercise. Second, the estimates of underlying seasonality from the demand system are more sensitive to the identification assumption employed in Einav (2007), that the unobservable component of the decay is independent of the choice of release date. Therefore, it may be useful to search for alternative sources of information about these parameters. Finally, a simple inspection of the results from the demand estimation suggests that distributors have a different seasonal pattern in mind when deciding about release dates compared to the seasonal pattern estimated. This calls for a more formal treatment, which would establish and quantify this pattern more analytically.

More generally, one may think about this application in the context of standard demand

FIGURE 6
Seasonal Effects in New Releases



Notes: This figure plots the estimated release effect, which is defined as the contribution of new movies to the competition effect (see Einav 2007 for more details). It is then averaged over the 15 years of the data. The dashed lines stand for two standard errors from the average (ignoring the implicit standard error that comes from coefficient uncertainty). The number of weekly dummies is 56 to account for the timing variation in U.S. holidays across the years (see Einav 2007 for details).

estimation. One strategy is to estimate a demand system combined with assumptions on the game played among firms. This approach is efficient, provided that the assumptions regarding firms' behavior are correct. It provides inconsistent estimates, however, if these assumptions are incorrect. This approach does not allow us to test for the optimality of firm behavior, as it is assumed. A second strategy is to estimate demand parameters from demand data alone, and then use these estimates to test the optimality of firms' behavior. This is the approach taken in this paper for two main reasons. First, it is computationally not feasible to estimate movie-specific qualities as parameters of the timing game; these qualities enter nonlinearly, requiring a numerical search procedure over many parameters. Second, the underlying demand parameters obtained from the demand system are quite different from industry wisdom, questioning the plausibility of pooling demand- and supply-side moments. Instead, I find it more informative to obtain these set of estimates separately and compare them.

Consequently, the seasonal estimates resulting from estimating the timing game should be thought of as the perceived underlying demand, that is, the underlying demand that rationalizes the observed release pattern. The interesting exercise is to compare these estimates to those derived from the demand system. As it turns out, these two sets of estimates of underlying demand are quite different. While I view this as some indication for bounded rationality of distributors (see later), it is perfectly consistent with the reverse interpretation: if one believes that the timing game is specified correctly and that distributors are fully rational, then the different patterns call into question the validity of the identification assumption used to obtain the demand estimates.

B. Specification

The general setup for the estimation is as follows. I choose several time windows ("seasons") within the year and take the set of movies that were released within the specified season as given. I then analyze the choice of the week within the season during which the movie is released. The motivation for this assumption comes from the prerelease timing data described in Section III. Distributors decide far in advance that a certain movie is scheduled for around, say, Memorial Day, but only later decide about the specific date on which it is released. Moreover, most changes of previously announced release dates do not shift the date by much.

The empirical model developed in Section II provides the framework for analysis. For the model to be taken to the data, I still need to specify the particular functions and the model parameters. In doing so, I am guided by two main considerations. First, the computational burden dictates a restricted choice of K (the number of weeks included in each season) and N (the number of strategic players), and a relatively small number of parameters. This is done by setting the length of each season to 5 weeks ($K = 5$), a choice which is guided by the switching behavior described in Figure 4. I let N be equal to 2–6, depending on the specification and the number of parameters estimated. I also estimate only a small number of parameters. As mentioned, the second important consideration for specifying the functional forms is my attempt to evaluate the timing decisions made by distributors.

I assume that each season represents an independent timing game. In each season, each of the N players (those that eventually released their movie during that season) chooses one of K weeks (that lie within the season) during which his movie is released. Movies are generally released on Fridays, consistent with analyzing the timing decision at the weekly level. To further reduce the dimensionality problem, I choose to model only the best N movies within the season as strategic players. The quality measure of each movie is given by the point estimate of the movie fixed-effect estimated in Einav (2007). I assume that these movies play against each other, conditioning on the observed release dates of all other movies.[22]

Given that all movies remain in the market for longer than one week, not only do the "active" players (top quality movies) condition on the release pattern of the lower quality movies within the season, but they also condition on the release pattern of all movies in adjacent seasons. While conditioning on the release dates of movies from the preceding season is sensible, it is questionable whether it is valid to assume that movies can condition on the release dates of movies in the subsequent season. I justify this assumption by the fast decay of box office revenues, which implies that the effect of movies that are released more than one week apart is relatively small, and hence has little effect on strategic considerations. I use these other movies and their observed release dates to calculate the counterfactual box office revenues.

For estimation, I choose four annual release seasons, which are all centered around a dominant release date. These are Presidents' Day, Memorial Day, Fourth of July, and Thanksgiving.[23] Each season includes the dominant week, and 2 weeks before and after, adding up to 5 weeks in each season (i.e.,$K = 5$, as specified earlier). Thus, I use a total of 60 seasons (four seasons over 15 years), on which the estimates are based. The number of movies in each season is between 6 and 17 with a mean of 11.2 and standard deviation of 2.34, but movie quality is very skewed. For example, the estimated quality of the top three movies accounts for 44–91% (with a mean of 66%), as a fraction of the total quality of all movies in the season. Thus, restricting attention to only the top movies accounts for the majority of the industry box office revenues in the season in which they are released.

As explained in the previous section, I keep the nested logit specification, which was the basis for the demand analysis in Einav (2007), and use the point estimates from the demand system, but free up some parameters. Specifically, I assume that the known portion of

22. A reasonable approach would be to assume that the order of moves is dictated by the ordering of the movie qualities, the biggest movie playing first. This implies that the smallest movies condition their decisions on the release dates of the bigger ones, but not vice versa (which is what I do in this paper). Given the computational restrictions, such an approach would rely on the decisions of the small, less strategic, players. For these movies, it is not clear that we need the "high-powered" structural game for estimation. Rather, given that they have no real strategic effect, we can estimate each movie's decision separately.

23. Christmas, a highly popular release date, is not used for the analysis for two reasons. First, the timing of many Christmas releases is driven by Academy Award eligibility requirements rather than by strategic motives. Second, unlike the other seasons, Christmas is not characterized by a single popular release date; the entire second half of December is popular among moviegoers.

TABLE 1
Estimation Results, Pooling All Seasons

A. Assuming better movie moves first						
Number of strategic movies (N)	**3**	**4**	**5**	**3**	**4**	**5**
η	1.13 (0.66)	1.14 (0.63)	1.12 (0.60)	1.18 (0.68)	1.20 (0.64)	1.18 (0.61)
α				1.72 (2.42)	2.12 (2.23)	2.03 (2.17)
Log likelihood	−283.3	−378.0	−472.9	−283.3	−377.9	−472.8
B. Assuming random (uniform) order of moves						
Number of strategic movies (N)	**3**	**4**	**5**	**3**	**4**	**5**
η	0.28 (1.13)	0.31 (1.18)	0.70 (1.38)	1.58 (1.24)	1.86 (1.38)	2.82 (1.70)
α				23.51 (16.52)	21.37 (14.40)	23.24 (16.31)
Log likelihood	−284.8	−379.8	−474.7	−284.0	−378.9	−473.3

Notes: The table presents the results from a set of specifications of the timing game. Panel A takes the order of moves as given, with the better movie moving first, while Panel B assumes a uniform distribution over all order permutations. Standard errors in parentheses. For comparison, note that the log likelihood of a fully random release date choice (i.e., $\eta = 0$) is $60 \cdot \ln(5^{-N})$.

distributors' profits takes the following form:

$$\hat{\pi}_j(r_j, r_{-j}; \gamma) = \sum_{t=r_j}^{r_j+H} \hat{s}_{jt}(r_j, r_{-j}; \alpha, \sigma) = \sum_{t=r_j}^{r_j+H} (\hat{D}_t^{\sigma} + \hat{D}_t)^{-1} \exp((\theta_j - \lambda(t-r_j) + \alpha\tau_t)/(1-\sigma)), \tag{13}$$

[Equation amended after online publication date September 29, 2009.]
where

$$\hat{D}_t = \sum_{k \in J_t(r_j, r_{-j})} \exp((\theta_k - \lambda(t - r_k) + \alpha\tau t)/(1 - \sigma)), \tag{14}$$

and θ_j is the estimated quality of the movie, λ is the estimated decay parameter, r_j is the (endogenous) movie's release decision, and τ_t is the estimated underlying demand. H is the length of the period that is taken into account by distributors when making their release decision. The choice of H is guided by computational limitations, so I choose $H = 2$, thereby restricting distributors to base their decisions on the first three weeks after the release. $J_t(r_j, r_{-j})$ is the set of movies that play on week t, which depends on the observed release dates of the nonstrategic movies as well as on the (endogenous) release decisions, r_j and r_{-j}, of the strategic movies which are being modeled.

That is, I use Equation (11) with small modifications. First, I assume that distributors make their decisions under the assumption that $\xi_{jt} = 0$ for any j and t.[24] Second, while all the parameters in the profit function are taken as given (based on the demand estimates), I introduce a new parameter, α. In the nested logit demand system, this parameter is restricted to be 1. Freeing it up allows distributors to overweight ($\alpha > 1$) or underweight ($\alpha < 1$) the estimated underlying demand.

Finally, an additional parameter to be estimated in all specifications is η, the precision of the logit error term as described in Section II. It does not show up in Equation (13) because it affects $\pi = \hat{\pi} + \varepsilon$ only through the error term, but not through $\hat{\pi}$. The results reported below also use different assumptions regarding the order of moves (see Section II).

C. Results

Table 1 present the estimation results for different choices of N, the number of strategic movies. Panel A presents results that are based

24. One could think of imputing ξ_{jt} from the demand system, and assuming that the ξ_{jt}'s are related to the movie-specific decay pattern. Doing so changes the results very little. This is because the variation of ξ_{jt}'s is very small and hence has little effect on the strategic considerations.

on an order of moves (of the sequential game) where the better movie moves first, while panel B present results where I allow a uniform distribution over all orders (permutations of the N players). Although the qualitative results (discussed below) are similar across both panels, the random order leads to somewhat less stable results and lower statistical significance of the coefficients, so I focus my discussion on the results of panel A, which constitutes my preferred specification.

Overall, the results are quite stable across different choices of N. In all specifications, the estimate of η is positive and significant at a 10% confidence level. Recall that η is the precision of the error term. Alternatively, one can also think of η as the parameter on the deterministic component of payoffs. An insignificant η would imply that the release date decisions appear random with respect to the modeled payoffs, and a negative η would imply that the modeled payoffs are negatively associated with the release date decision. Therefore, the positive and (marginally) significant estimate of η suggests that the model for payoffs, together with the estimated demand parameters, is indeed useful in explaining the release date decisions.

Perhaps more interesting is the estimate of α. Across all specification, the point estimates of α are consistently above 1, which is the implied value of the nested logit demand system.[25] This suggests that movie distributors overweight underlying seasonality (relative to competition from other movies) when they make their release date decision. In other words, to best rationalize the observed release date decision, the estimated underlying demand estimates need to be about doubled; that is, the spike of underlying demand in, say, Memorial Day weekend needs to be twice as large to rationalize the clustering of hit movies released on that weekend.

Thus, the results taken together suggest that although distributors tend to respond to underlying demand and to competition from other movies, as implied by the demand model, they appear to be too clustered in holiday weekends. To make this statement more precise, and to provide more interpretable figures, I construct a measure for clustering. In a given season, for a given choice of N, I define the clustering measure as the average fraction of quality released on the holiday weekend. Let θ^i_m be the quality of movie i, which is released in season m, so the average clustering measure across M markets is given by

$$\text{clustering} = \frac{1}{M}\sum_{m=1}^{m=M}\left(\sum_{r_i \text{ is holiday}} \theta^i_m\right)\Big/\left(\sum_i \theta^i_m\right). \tag{15}$$

This is the actual clustering measure. I construct the corresponding counterfactual by using the expected clustering measure, where the expectation is taken over the idiosyncratic noise in the empirical model, and over the distribution of the permutations of order of play. One should note that the clustering measure is between 0 and 1, and that with $K = 5$, a random assignment of movies into release dates yields a clustering measure of $K^{-1} = 0.2$.

Across all seasons and years, the average clustering measure ranges from 0.35 to 0.37 (across choices of N) compared to the random assignment measure of 0.2. Computing expected clustering measures using the point estimates from the various specifications of Table 1 that impose a value of $\alpha = 1$, I obtain measures that range from 0.16 to 0.25. That is, the nested logit demand model implies a much more even distribution of movie quality within a season, requiring overweighting of underlying seasonality to rationalize the much higher actual clustering observed in the data.

Table 2 presents some results when each release season is estimated separately. I am hesitant to experiment much with various specifications because each such specification relies on only 15 independent seasons. Table 2 is still instructive in showing that pooling together all seasons may hide important heterogeneity. Specifically, the results presented in Table 2 show that the model does a pretty good job in explaining release date decisions in the summer (i.e., around Memorial Day and Fourth of July), while the release date decisions around Thanksgiving, and even more so around Presidents' Day, are hardly associated with expected profits, as modeled and estimated by the demand system.

D. Discussion

The main conclusion from the empirical analysis of this paper is that distributors seem

25. None of the estimates of α is significantly different from 1 (or from 0) at reasonable confidence levels. This may not be surprising given the small number of independent seasons (60) used for estimation, which makes standard errors large. However, given the fairly stable estimate of α across choices of N, interpreting and discussing the point estimates may not be unreasonable.

TABLE 2
Estimation Results, by Season

Number of strategic movies (N)	**2**	**3**	**4**	**5**	**6**
Presidents' Day					
η	0.00 (0.01)	0.08 (1.41)	0.07 (1.36)	0.00 (0.0)	0.00 (0.01)
Log likelihood	−45.1	−67.6	−90.1	−112.7	−135.2
Memorial Day					
η	2.20 (1.25)	2.08 (1.18)	1.79 (1.19)	1.90 (1.16)	1.98 (1.19)
Log likelihood	−46.6	−70.8	−95.3	−119.1	−143.2
Fourth of July					
η	0.00 (0.04)	1.87 (1.80)	1.94 (1.66)	2.34 (1.46)	1.74 (1.40)
Log likelihood	−48.3	−71.8	−95.7	−119.1	−143.9
Thanksgiving					
η	0.23 (1.41)	0.49 (1.17)	0.83 (1.05)	0.56 (1.04)	0.55 (0.97)
Log likelihood	−48.3	−72.3	−96.2	−120.6	−144.7

Notes: The table presents the results from a set of specifications of the timing game, separately by season. All results take the order of moves as given, with the better movie moving first (as in panel A of Table 1). Standard errors in parentheses. For comparison, note that the log likelihood of a fully random release date choice (i.e., $\eta = 0$) is $15 \cdot \ln(5^{-N})$. Note that almost no estimate is statistically significant at reasonable confidence levels. This is mainly due to the small number of observations once each season is allowed to have its own parameters.

to cluster their release dates more than they should. While the spirit of the results is somewhat similar to the conclusions drawn in Einav (2007), the results in the current paper are largely driven by different variation in the data. Einav (2007) uses cross-seasonal variation and "price-taking" assumptions, while here I use within-season variation and allow for strategic effects. Thus, at least a priori, there was no reason to assume that the two analyses would yield similar qualitative results. The fact that they do, therefore, strengthens these conclusions. In what follows, I propose several explanations that may help understand these findings.

The first line of interpretations is consistent with the assumptions of optimizing behavior by studios. There are several industry features, not incorporated in my model, that could cause more movie clustering around high-demand weekends than predicted by the model. First, uncertainty may play an important role. While a complete information equilibrium may have the movies spreading out over the different weeks, uncertainty may result in more movies being released on the better weeks. This over-clustering may prove inefficient ex-post, but may be optimal ex-ante. To gain intuition for why this may be the case, consider a two-by-two game in which the ex-post profits are given by

	Holiday	Non-Holiday
Holiday	100, 0	100, 40
Non-Holiday	40, 100	10, 0

where the motivating story is that the market size is 100 and 40 in the holiday and non-holiday weeks, respectively. Movie 1 (the row player) always obtains the whole market, independently of competition, while movie 2 makes positive revenues only if movie 1 is not present. The equilibrium of this model is for movie 1 to release on the holiday weekend, and for movie 2 to release on the non-holiday weekend, that is, no clustering. Consider now an extreme uncertainty, in which both movies are identical ex-ante, and with probability 0.5 each movie wins the whole market if the two compete head-to-head. The new (ex-ante) payoff matrix would be as follows:

	Holiday	Non-Holiday
Holiday	50, 50	100, 40
Non-Holiday	40, 100	10, 20

yielding a unique Nash equilibrium in which both movies are released on the holiday weekend, that is, clustering. Trying to allow uncertainty in the empirical model is conceptually possible but computationally very intensive.

One can also rationalize the overclustering result by having the true value of a holiday release being greater than what we think it is. For example, this may be due to repeated game effects (Chisholm 1999), which may change the static "value" of releasing on a certain week. If a distributor who releases a movie on, say, the Fourth of July is more likely to capture the same week in future years, a Fourth of July release may be more attractive than it is estimated to be. Alternatively, if there are nonmonetary benefits (e.g., prestige) to the distributor (or to the director and the actors) from releasing on holiday weekends, this may also generate overclustering effects.[26] Third, as is well known, the nested logit specification used to obtain the demand estimates assumes that all movies are equally good substitutes of each other, proportional to their market shares. If the top movies every season are of different genres, the revenues of these movies may be less affected by clustering with other movies.

A very different line of explanation is a behavioral one. Movie distributors could be overconfident as to the relative quality of their movie (Camerer and Lovallo 1999) or could simply err in their assessment of the underlying demand in the industry. After all, the "learning from experience" argument, according to which economic agents cannot err for a long period, may not work in the motion picture industry. To learn from experience, distributors have to first obtain enough experience and then be able to use it properly. In particular, the information about the seasonal pattern comes only once a year, and the high uncertainty about movie quality makes inference difficult regarding the separation between the underlying demand and the movie quality effects. With each movie having its own identity, a controlled experiment of releasing the same (or very similar) movie in different dates is not feasible. For decades, the industry has followed the same release pattern, according to which big hits are released on big weekends. Thus, there are no natural experiments that make it easy to distinguish between higher movie quality and higher underlying demand. Any deviation from the "predicted" seasonal pattern (for example, successful movies in October) is typically interpreted by industry observers as an extremely good movie in the wrong season rather than as a decent movie in a mediocre season. In other words, there is very little bad feedback after a bad release decision.

Even if distributors are fully rational, conservatism may lead them to stick to the traditional release pattern. This conservatism may be magnified if we think of the institutional context and the potential agency costs in the industry. Top directors and actors do not want to see their films fail because of a poor marketing decision. Thus, considering the traditional release pattern in the industry, they frequently lobby for a traditionally good choice of release dates. Distributors are likely to be conservative and satisfy these requests, rather than risk their jobs, reputation, and future business. By sticking to the traditional release pattern they can be adequately evaluated by the market.[27] This is not the only example where the motion picture industry seems to be conservative and to follow tradition. Other examples include the current uniform ticket pricing policy in the industry (Orbach and Einav 2007), the use of stars in the industry (Ravid 1999), or the massive capacity expansion that took place in the 1990s and has recently led many of the largest chains of movie theaters to file for bankruptcy (Davis 2006a).

Finally, it should be noted that in recent years distributors started to experiment more with less traditional release decisions. After relative successful early May openings of *Gladiator* and *The Mummy Returns* in 2000 and 2001, the distributors of *Spiderman*—an anticipated blockbuster much before its actual release—decided to release it on May 3, 2002. Ten years earlier such a move would have been unheard of.

V. CONCLUSIONS

This paper develops a new empirical model of discrete games to study the release date timing game played by movie distributors. The timing game is formulated as a sequential game with private information, with distributors choosing among a small set of release weekends. The main empirical finding is that movie distributors overcluster their release dates, with too many good movies released on big holiday weekends. As a whole, the results complement and strengthen similar (though weaker) conclusions found in Einav (2007).

26. Consistent with this idea is the sentence I often heard while interviewing industry executives: "Economics? This industry is not about economics; it is all about egos. . . ."

27. This is in the spirit of "you cannot be fired for buying IBM." For a formal treatment, see Zweibel (1995).

In analyzing the timing game, the assumption is that by the time the release date is chosen, many of the characteristics of the movie and of competing movies are already set, generating heterogeneity among competitors. This provides each player with a payoff-relevant identity. Accounting for this feature using existing models of discrete games is difficult. This paper therefore proposes a new estimable game structure that tries to relax this limitation. The model specifies a sequential game structure with asymmetric information. The game has a unique equilibrium that can be solved using a simple algorithm. This allows estimating a game even in the absence of symmetry restrictions on the payoff structure.

This last feature is crucial for a large set of applications. In particular, such flexibility would be a necessary property of an empirical model of discrete games that attempts to model location choice together with a state-of-the-art model of price-setting behavior. As the demand literature in industrial organization keeps going in the direction of more flexible substitution patterns, one has to use product choice models that allow for a more flexible functional form for payoffs. Combining these two literatures is an important direction for future research. The empirical model proposed in this paper is a way that may facilitate such work.

REFERENCES

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

Andrews, D., S. Berry, and P. Jia. *"Confidence Regions for Parameters in Discrete Games with Multiple Equilibria, with an Application to Discount Store Locations."* Manuscript, Yale University, 2007.

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

Bajari, P., H. Hong, and S. Ryan. *"Identification and Estimation of a Discrete Game of Complete Information."* Manuscript, University of Minnesota, 2008.

Beresteanu, A., I. Molchanov, and F. Molinari. *"Sharp Identification Regions in Games."* Manuscript, Duke University, 2008.

Berry, S. "Estimation of a Model of Entry in the Airline Industry." *Econometrica*, 60, 1992, 889–917.

Berry, S., J. Levinsohn, and A. Pakes. "Automobile Prices in Market Equilibrium." *Econometrica*, 63, 1995, 841–90.

Berry, S., and J. Waldfogel. "Free Entry and Social Inefficiency in Radio Broadcasting." *RAND Journal of Economics*, 30, 1999, 397–420.

Borenstein, S., and J. Netz. "Why Do All the Flights Leave at 8 a.m.? Competition and Departure-Time Differentiation in Airline Markets." *International Journal of Industrial Organization*, 17, 1999, 611–40.

Bresnahan, T., and P. C. Reiss. "Entry into Monopoly Markets." *Review of Economic Studies*, 57, 1990, 531–53.

Camerer, C., and D. Lovallo. "Overconfidence and Excess Entry: An Experimental Approach." *American Economic Review*, 89, 1999, 306–18.

Caruana, G., and L. Einav. "A Theory of Endogenous Commitment." *Review of Economic Studies*, 75, 2008, 99–116.

Chisholm, D. C. *"The War of Attrition and Optimal Timing of Motion-Picture Releases."* Manuscript, Suffolk University, 1999.

Ciliberto, F., and E. Tamer. *"Market Structure and Multiple Equilibria in Airline Markets."* Manuscript, Northwestern University, 2008.

Corts, K. S. "The Strategic Effects of Vertical Market Structure: Common Agency and Divisionalization in the U.S. Motion Picture Industry." *Journal of Economics and Management Strategy*, 10, 2001, 509–28.

Davis, P. J. "Measuring the Business Stealing, Cannibalization, and Market Expansion Effects of Entry in the U.S. Motion Picture Exhibition Market." *Journal of Industrial Economics*, 54, 2006a, 293–321.

Davis, P. J. "Estimation of Quantity Games in the Presence of Indivisibilities and Heterogeneous Firms." *Journal of Econometrics*, 134, 2006b, 187–214.

Draganska, M., S. Misra, V. Aguirregabiria, P. Bajari, L. Einav, P. Ellickson, D. Horsky, et al. "Discrete Choice Models of Firms' Strategic Decisions." *Marketing Letters*, 19, 2008, 399–416.

Einav, L. "Seasonality and Competition in Time: An Empirical Analysis of Release Date Decisions in the U.S. Motion Picture Industry." Ph.D. dissertation, Harvard University, 2002.

Einav, L. "Seasonality in the U.S. Motion Picture Industry." *RAND Journal of Economics*, 38, 2007, 127–45.

Einav, L., and A. Nevo. "Empirical Models of Imperfect Competition: A Discussion," in *Advances in Economics and Econometrics: Theory and Applications*, Vol. II, Chapter 3, edited by R. Blundell, W. K. Newey, and T. Persson. Ninth World Congress. Cambridge: Cambridge University Press, 2006, 86–96.

Einav, L., and A. Ravid. *"Stock Market Response to Changes in Movies' Opening Dates."* Manuscript, Stanford University, 2007.

Ellickson, P. B., and S. Misra. *"Enriching Interactions: Incorporating Revenue and Cost Data into Static Discrete Games."* Manuscript, University of Rochester, 2007.

Goettler, R. L., and R. Shachar. "Spatial Competition in the Network Television Industry." *RAND Journal of Economics*, 32, 2001, 624–56.

Lukk, T. *Movie Marketing: Opening the Picture and Giving It Legs*. Los Angeles: Silman-James Press, 1997.

Mazzeo, M. J. "Product Choice and Oligopoly Market Structure." *RAND Journal of Economics*, 33, 2002, 1–22.

McKelvey, R. D., and T. R. Palfrey. "Quantal Response Equilibria in Extensive Form Games." *Experimental Economics*, 1, 1998, 1–41.

Moretti, E. *"Social Learning and Peer Effects in Consumption: Evidence from Movie Sales."* NBER Working Paper No. 13832, 2007.

Moul, C. C. "Measuring Word of Mouth's Impact on Theatrical Movie Admissions." *Journal of Economics and Management Strategy*, 16, 2007, 859–92.

Nevo, A. "Measuring Market Power in The Ready-to-Eat Cereal Industry." *Econometrica*, 69, 2001, 307–42.

Orbach, B. Y., and L. Einav. "Uniform Prices for Differentiated Goods: The Case of the Movie-Theater Industry." *International Review of Law and Economics*, 27, 2007, 129–53.

Orhun, Y. *"Spatial Differentiation in the Supermarket Industry."* Manuscript, UC Berkeley, 2005.

Pakes, A., M. Ostrovsky, and S. Berry. "Simple Estimators for the Parameters of Discrete Dynamic Games, with Entry/Exit Examples." *RAND Journal of Economics*, 38, 2007, 373–99.

Quint, D., and L. Einav. "Efficient Entry." *Economics Letters*, 88, 2005, 278–83.

Ravid, S. A. "Information, Blockbusters, and Stars: A Study of the Film Industry." *Journal of Business*, 72, 1999, 463–92.

Reiss, P. C. "Empirical Models of Discrete Strategic Choices." *American Economic Review Papers and Proceedings*, 86, 1996, 421–26.

Seim, K. "An Empirical Model of Firm Entry with Endogenous Product-Type Choices." *RAND Journal of Economics*, 37, 2006, 619–40.

Simonsohn, U. *"eBay's Crowded Evenings: Competition Neglect in the Field."* Manuscript, UC San Diego, 2008.

Sweeting, A. *"The Strategic Timing of Radio Commercials: An Empirical Analysis Using Multiple Equilibria."* Manuscript, Duke University, 2008.

Tamer, E. "Incomplete Simultaneous Discrete Response Model with Multiple Equilibria." *Review of Economic Studies*, 70, 2003, 147–65.

Toivanen, O., and M. Waterson. "Market Structure and Entry: Where's the Beef?" *RAND Journal of Economics*, 36, 2005, 680–99.

Watson, R. "Product Variety and Competition in the Retail Market for Eyeglasses." *Journal of Industrial Economics*, 57(2), June 2009, 217–251.

Zweibel, J. "Corporate Conservatism and Relative Compensation." *Journal of Political Economy*, 103, 1995, 1–25.