Competitive Dynamics of Genre Choices in the Film Production

Ping Zheng*
Economics Department
Indiana University
Bloomington, IN 47405
pzheng@indiana.edu

Draft: April 20

* Ping Zheng is a third-year doctoral student at the Economics Department of Indiana University Bloomington. This paper is composed to fulfill the third-year paper requirement as part of the program study. It also shows the author’s research interest. The current draft is incomplete as it does not include the part of empirical estimation.
Competitive Dynamics of Genre Choices in the Film Production

Abstract

This paper considers an industry dynamic discrete-choice model (the DP model thereafter) to characterize the strategic behaviors of major film makers in the production choices of film genres in the U.S. motion picture industry. Given the fact that film production decisions are farsighted, the DP model is better in the sense that it not only models the players’ decisions based on the current state status, but also based on the players’ expectations of the decision probabilities from their opponents integrated over all the possible future states. A further empirical work is needed to estimate the model parameters and test the model’s goodness of fit.

Keywords: Film Genres, Film Production, Motion Picture Industry, Dynamic Discrete Choice Games.
1. Introduction

The U.S. motion picture industry is a multi-billion dollar business. Spending only on theatrical tickets was over $9 billion in the U.S. (and Canada) and around $17 billion outside the U.S. in 2007 (Motion Picture Association of America, www.mpaa.org), let alone the tremendous revenues from ancillary markets (DVD rental and sales and merchandising). The motion picture industry employs over half a million people in the U.S. (U.S. Department of Labor 2007) It is also currently the number one export market for the U.S. From these statistical facts we see that this industry has a high economic importance in the global economy. Moreover, this industry has high cultural significance. Hollywood, the representation of American motion picture industry, has earned the worldwide recognition as one of the two outstanding cultural symbols (the other being Lady Liberty) of America, which has inspired millions of young Americans as well as people around the world to pursue the glorious American dream.

Due to both economical and cultural significance the amount of academic research on issues related to the motion picture industry has risen sharply. Most of the studies focus on understanding the determinants, from both supply and demand sides, of consumer demand for movies and forecasting film theatrical performance (box office revenues). While these studies in general ignore the competitive nature of this industry, other studies that account for strategic interactions among films competing for box office revenues mainly take place on the arena of film releasing time games. An area that is overlooked yet as important is the strategic behaviors of film makers who supervise the material and quality of films that in turn to a great extend determine the success or failure of a film. Thus, in this study I look into the strategic competition that takes place in the film production channel. Specifically, I try to capture how competitive film makers decide which film genres to produce each period and how their decisions in the
previous periods would affect the late periods. To examine the performance of my model, I will compare the model’s out-sample fit with the actual frequencies observed in the data.

I focus only on the major (six) studios in the U.S. since for one they represent the majority of the motion picture industry and for two there is a systematic difference between independent film production companies and studios. Since (major) film studios have their own capital and/or well-established financing channels, that is to say film financing is not a big concern to the studios, most of their attention is drawn on the expected future revenue of a film, which whether or not would recover the production costs and would there be a room to spend on marketing.

Since a studio’s production decision is farsighted, I employ a dynamic programming model to address the issues that are related to one particular production activity of film studios, the genre choice. I model a studio’s decision on which genre of a film to produce each period as a multinomial-choice game. In each period the studios simultaneously decide whether nor not to invest, and if so which genre of a film to produce given the current genre size (the number of films in each genre recently put into production, as these films are potential competitors for the future revenues) and their beliefs on the choices of their opponents.

Film production is certainly far complicated than choosing genres. The production activities for a studio include three phrases, pre-production, such as “green-lighting” process and preparation, production that is the actual shooting stage, and post-production, such as film editing. Even though I assume financing do not impose a constraint on a major studio, these other activities, such as how much to invest in each film, all involve strategic

---

2 One important difference is that independent film makers are under much greater financial pressure than the major film studios. As a result, the strategic behaviors of independent film makers must also be subject to this constraint, which I don’t take into account here.

3 This is an industrial terminology that refers to the process that the initial decision to approve or decline a project.
interactions both within and among studios, explained as studios competing for limited resources (talents). Thus, completely modeling film production is a very challenging job. I take modeling the genre choice made by studios as the initial step in this field of my interest. I hope to look into more complicated issues in my future research.

The rest of paper is organized as follows. In section 2 of literature review I first briefly go through some of the research that has done in the U.S. motion picture industry, then I provide some background on the dynamic programming model and the recently development on its estimation methods. Section 3 includes the framework of game, the detailed modeling as well as the estimation method.

2. Literature review

2.1. Literature on the motion picture industry

There is a rich stream of marketing and relatively fewer economic literature on modeling the motion picture industry. A majority of these studies focus on understanding the determinants of consumer demand for movies and forecasting films’ theatrical performance (box office revenues) without explicitly accounting for competition from other films. Such studies include consumer purchase intentions (Sawhney and Eliashberg 1996) from the demand side, advertising (Zufryden 2000) and number of screens and distribution strategy (Elberse and Eliashberg 2003) from the supply (distributor) side, as well as various film attributes such as star powers (Alberta 1998), clustering (Jedid, Krider and Weinberg 1998), seasonality (Einav 2007), ratings (Vany and Walls 2002) and critical reviews (Eliashberg and Shugan 1997). Also note that most of these empirical work has uniformly employed a regression type of analysis.
There are a smaller number studies that have accounted for the role of competition in films’ box office revenues. Of them, Zhang and Kadiyali (2006) is the only paper so far that has adopted a dynamic structural modeling technique, dynamic programming, to look into studios’ strategic decisions on films’ release date pre-announcement activities. Specifically, the authors model major Hollywood movies’ simultaneous choices of release dates as a dynamic market entry/exit game, where the market is defined as any five-week holiday season (centered at a particular holiday) over certain periods. The authors show that their model outperforms a number of choices of static structural models in terms of both in-sample fit and out-sample forecasting. My work is similar to Zhang and Kadiyali (2006) as we both characterize our competitive dynamics, yet in different channels, as a simple discrete choice (market entry/exit) game.

2.2. Related literature on competitive dynamics

The equilibrium notion under an industry dynamic framework is so-called Markov Perfect Equilibrium (MPE) proposed by Ericson and Pakes (1995). It basically states that at equilibrium firms choose the optimal strategies (investment, entry and exit decisions) by maximizing their present discounted value given expectations about the evolution of their competition. Furthermore, at equilibrium those expectations are fully consistent with the process generated by the optimal decisions of all firms. The authors show the existence of such a (rational expectations) equilibrium with a finite number of heterogeneous agents subject to idiosyncratic shocks.

Two issues in general arise in the dynamic programming (DP) model. One is the existence of multiple equilibria that is a prevalent feature in most empirical games where best response functions are nonlinear in other players’ actions. To solve this problem, the common approach is to impose restrictions that guarantee equilibrium uniqueness. The other issue is
computational costs in estimation of the model that involves solving fixed points of the system of best response functions nested inside an optimization problem (maximizing the log likelihood). This feature is even typical in dynamic discrete choice games, known as “curse of dimensionality” in the sense that the cost of computing an equilibrium increases exponentially with the number of players. These two issues have limited the scope of applications of the DP model. Fortunately, recent developments in estimation methods have made the DP model possible to be extended to cover a wide range of applications. In dynamic discrete choice games, for example, Aguirregabiria and Mira (2007) proposes a class of pseudo maximum likelihood (PML) estimators that are shown to solve the above two issues under certain regularity conditions. PLM has also been tested by a variety of empirical work and has been shown to work well in applications. Thus in this study I will employ this estimation method, whose details will be discussed along the pass estimating my model.

3. Modeling and estimation

3.1. The game

The production decision of a film studio is constantly repeated. That is, a film studio first starts with a film of particular genre. In the process of its development, the studio discovers a new idea and/or makes new deals about another film, which if is approved will become the studio’s new project. This process repeatedly takes place along with the development of the existing projects. I observed in the data that some studios start new projects sometimes as frequent as every week, and other times there are a couple of months apart between the start dates of two projects; and other studios devote less into the production, producing only three or four films during a given year.
This discrepancy among studios’ production behaviors coincides with firms’ profit maximization theory. As a studio plays two roles at the same time, one as a film production company and the other as a film distributor, it is reasonable that a studio shifts its focus from one channel to the other based on its expectation of the overall profitability from time to time. A typical example is 20th Century Fox. This studio produced more films, on average ten films per year, and distributed less, about three/four films per year, in the 80s and early 90s. Then, it started to shift its attention from the production business to distribution. Nowadays, the number of films produced and distributed (excluding the one that are produced by the studio) each year goes the other way around. Thus, when modeling film studios’ production behaviors it is important to also take into account the competition between the two channels (production and distribution) within a studio. However, due to the tractability in this paper I ignore this factor, which could possibly be a future research area.

The game is of infinite horizon. There are fixed number (six) of major film studios that compete with each other for the box office receipts (theatrical sales). Each period studios simultaneously make decisions on which genre of a film to be developed, based on one market condition (or state), that is, the current genre size (the number of films in each genre recently put into production). This condition forms the competition set that a studio could foresee in future. If there already exist a number of films of one particular genre out there in production and if a studio expects his opponents are likely to invest in this genre, then the studio would probably postpone its decision on investing in the same genre at current period and take another alternative due to the possible crowding-out effect in future. However it does not seem to be entirely true. By closely examining the data, I observe that even though the genre choices by studios are diversified each period, there is some time when studios heavily invested in one
particular genre (classified as the dominant genre; see Table 1. in the data section). This phenomenon can be explained by the seasonality nature of film, as films of the same genre tend to be targeted at the same holiday season.

There are certainly a lot of factors that contribute to a studio’s decision on which film to produce each period. Sometimes it is simply due to the appeal of a story; sometimes it could be the deals that a studio has made with famous directors and/or actors (Ravid 1999). These are the private information that is not observable to researchers. Yet there are some other observables that also influence a studio’s preference. One of such factors is a genre’s recent average performance, which somewhat indicates the trend of market demand over time. It is reasonable to assume that a studio likely choose a genre that has the highest average box office receipts in the recent history. This is supported by the evidence that studios tend to concentrate on a few particular genres among all classified genres (There are fifteen classified genres according to The Motion Picture Guide (Nash and Ross 1985), and eighteen genres according to Internet Movie Database (www.imdb.com)), and some genres (eg. western films) have almost disappeared from the big screen. In the current model setting, I do not consider this variable, a genre’s average box office receipts, as a pertinent state variable because in dynamic discrete-choice model state variables are required to have a discrete and finite support (Aguirregabiria and Mira 2007). Thus, for modeling purpose I include all the unobservables and the “inappropriate” observables into a stochastic error term.

The equilibrium, which I assume to be symmetric, is such that each period a studio makes his optimal decision on the genre choice of a film to produce or not investing in any film at all given the current state status, the genre size and his expectations on the genre choice of his opponents, to maximize his expected present value of the film’s box office revenues over all the
possible future states whose evolution (from period to period) is governed by a set of stationary transition probabilities. Following the notion of Ericson and Pakes (1995), I refer to this equilibrium as the Markov Perfect Nash Equilibrium (MPE).

3.2. The model

The model setup follows and extends the general framework of a dynamic discrete game with infinite time horizon described in Aguirregabiria and Mira (2007). In the current model, the time frequency used is month, denoted by \( t = 1,2, \ldots \infty \). The players are the major movie studios, whose number is fixed each period and denoted by \( N \). In time \( t \), a studio \( i \) observes two vectors of current states \( s_t = (x_t, \epsilon_{it}) \), where a deterministic component \( x_t \equiv a_t-1 \) is the genre choices of studio \( i \) and his opponents \( i \) during the previous period \( t - 1 \), and is common knowledge to all studios; \( \epsilon_{it} \) is the random component that captures the unobserved characteristics of studio \( i \) (or interpreted as the private shock). A studio \( i \)'s decision at time \( t \) is represented by \( a_{it} \in A \equiv \{0 \text{ or not invest}; \ g \in \{1,2, \ldots, 8 \} \} \), where \( A \) is the common choice set and \( g \in \{1,2, \ldots, 8\} \) stands for a particular genre choice from eight identified alternatives over certain period of time.

In each period, a studio chooses an action \( a_{it}^* \) that maximizes his expected payoffs, the discounted sum of each period’s expected payoffs over the entire time horizon. This is characterized by the recursive Bellman equation:

\[
V_{lt}(x_t, a_{lt}) = \pi_{lt}(x_t, a_{lt}) + \epsilon_{lt}(a_{lt}) + \gamma \sum_{x_{t+1}} V_{l,t+1}(x_{t+1})f(x_{t+1} | x_t, a_{lt}) \tag{1}
\]

where \( \gamma \) is the discounted factor. \( \pi_{lt}(x_t, a_{lt}) + \epsilon_{lt}(a_{lt}) \) is the current (expected) payoff, and \( \gamma \sum_{x_{t+1}} V_{l,t+1}(x_{t+1})f(x_{t+1} | x_t, a_{lt}) \) is thus the discounted weighted sum of the expected

---

4 Aguirregabiria and Mira (2007) uses a simple example of dynamic estimation of market entry/exit decision (i.e. a binary choice model), while here I extends to model a multinomial outcome model.
future payoff, weighted by the possible future states, where \( f(x_{t+1}|x_t, a_{it}) \) is the transition probability that governs the evolution of the state variable from the current state \( x_t \) to the future states \( x_{t+1} \) conditional on studio \( i \)'s current decision \( a_{it} \).

Since the state variable is the set of genre choices of all studios from last period, under the assumption of rational expectation \( f(x_{t+1}|x_t, a_{it}) \) can be conveniently estimated as the product of choice probabilities of all \( i \)'s opponents (from studio \( i \)'s point of view) given \( i \)'s current choice \( \prod_{j \neq i} P(a_{jt}|x_t) I(a_{it} = a_i) \). Not that studio \( i \)'s choice probability is the best response of the joint choice probabilities of all other studios.

The current (expected) payoffs consist of two parts, a “deterministic” part \( \pi_{it}(x_t, a_{it}) \) and a random part \( \epsilon_{it}(a_{it}) \), that are additively separable. This is one of the primitive assumptions that are standard in the literature for the purpose of tractability. In each period, the deterministic payoffs are a function of a studio’s expected revenue from developing a film of a particular genre, the production budget of that film and the number of films of the same and different genres in production from last period:

\[
\pi_{it}(x_t, a_{it}) = I(a_{it} > 0)[\beta_1 \log \bar{D}_{gt} + \beta_2 \log Q_{gt} + \beta_3 n_{sg} + \beta_4 n_{dg} + \beta_5 n_{sg}^2]
\]  

where \( I(\cdot) \) is an indicator function that takes value 1 if a studio decides to produce a film at time \( t \); otherwise, zero. \( \log \bar{D}_{gt} \) is the log of the anticipated demand for a film of genre \( g \), (to be estimated using static model) and I conjecture the sign of \( \beta_1 \) being positive. \( \log Q_{gt} \) is the log of production budget of the film, whose coefficient \( \beta_2 \) is expected to be negative. \( n_{sg} \) and \( n_{dg} \) are the number of films in the same and different genres recently put into production respectively. The effect of the number of films in the same genre is a mixture as I have described in the game subsection. Yet I conjecture that the sign of \( \beta_4 \) being positive as a few competitions would stimulate audience’s incentives of going to theatre, while the sign \( \beta_5 \) being negative as too many
films in the same genre form a large potential competition set and would have a crowding-out effect in future. On the other hand, films in different genres would not have or have less (if there is any) crowding-out effect so that $\beta_5$ is expected to be positive. Define $\beta \equiv (\beta_1, \beta_2, \beta_3, \beta_4, \beta_5)$ as the vector of payoff parameters. Finally, I normalize the (deterministic) payoffs for non-investors as a single parameter that is studio specific and time-invariant, that is, $\pi_{it}(x_t, 0) = \mu_i$. Define $\mu \equiv (\mu_1, ..., \mu_e)$ as the vector of non-investors’ payoff parameters.

Given the assumption of conditional independence, proposed by Rust (1987), of the error term $\varepsilon_{it}$ and the state variable $x_t$ across all players (studios) and over all periods, and assuming $\varepsilon_{it}$ follows the i.i.d. Type I Extreme Value distribution (McFadden 1974), it can be shown that the conditional choice probability of studio $i$ follows a logit form:

$$P(a_{it}|x) = \frac{\exp\{\pi_{it}(x_t, a_i) + \gamma \sum_{x' \in X} V_i(x') f(x' | x, a_i)\}}{\sum_{a_j \in A} \exp\{\pi_{jt}(x_t, a_j) + \gamma \sum_{x' \in X} V_j(x') f(x' | x, a_j)\}}$$

(3)

Let $P^*$ be an equilibrium, and let $V_i^{P^*}$ be the value function associated with the equilibrium. The Bellman equation (1) becomes:

$$V_i^{P^*}(x) = \sum_{a_i \in A} P_i^*(a_i | x) [\pi_i^{P^*}(a_i, x) + e_i^{P^*}(a_i, x)] + \beta \sum_{x' \in X} V_i^{P^*}(x') f^{P^*}(x' | x, a_i)$$

(4)

where $e_i^{P^*}(a_i, x) \equiv E(\varepsilon_i(a_i) | x, \sigma_i^*(x, \varepsilon_i) = a_i)$. Under the assumption of error distribution (i.i.d. Type I Extreme Value) and let $\sigma$ be the dispersion parameter, I thus obtain $e_i^{P^*}(a_i, x) = \text{Euler’s constant} - \sigma \ln(P_i^*(a_i | x))$.

Then the optimal decision $a_{it}^*$ is obtained by solving:

$$a_{it}^*(x_t) = \arg \max_{a_{it}} V_i^{P^*}$$

(5)
3.3. Estimation

Taking the equilibrium probabilities as given, equation (4) describes the vector of values \( V_{it}^P \) as the solution of a system of linear equations, and in vector form,

\[
(I - \beta F^P) V_{it}^P = \sum_{a_i \in A} P_i^P(a_i) \left[ \pi_i^P(a_i, x) + e_i^P(a_i, x) \right]
\]  

(6)

where \( F^P \) is a matrix with transition probabilities \( f^P(x'|x) \). Given that \( (I - \beta F^P) \) is of full rank, \( V_{it}^P = V(x, P^*) = (I - \beta F^P)^{-1}\{\sum_{a_i \in A} P_i^P(a_i) \left[ \pi_i^P(a_i, x) + e_i^P(a_i) \right]\} \). Substituting this solution of value function into equation (3), the equilibrium conditional choice probability can be characterized as a fixed point of a mapping, that is \( P_i^* \equiv \Psi_i(a_i|x; P^*) \), where

\[
\Psi_i(a_i|x; P^*) = \frac{\exp\left\{ \pi_{it}(x_t, a_i) + \gamma \sum_{x' \in X} V_i(x', P) f(x'|x, a_i) \right\}}{\sum_{a_j \in A} \exp\left\{ \pi_{jt}(x_t, a_j) + \gamma \sum_{x' \in X} V_j(x', P) f(x'|x, a_j) \right\}}
\]  

(7)

Define the set of structural parameters to be estimated as \( \theta \equiv (\beta, \mu, \sigma) \). The log likelihood function is given by:

\[
\ln(\theta) = \sum_{i=1}^{N} \sum_{t=1}^{T} I\{a_{it} = a\} \ln \Psi_i(a_i|x; P, \theta)
\]  

(8)

To estimate equation (6), I employ the Nested Pseudo Likelihood (NPL) method, proposed by Aguirregabiria and Mira (2007). The NPL algorithm describes as follows: Let \( \hat{\theta}_0 \) be an initial guess of the vector of studios’ choice probabilities. Given \( \hat{\theta}_0 \), the two-step Pseudo Maximum Likelihood estimator is given by \( \hat{\theta}_{2s}^{(1)} \equiv \arg \max_{\theta \in \Theta} \ln(\theta, \hat{\theta}_0) \). Given \( \hat{\theta}_0 \) and \( \hat{\theta}_{2s}^{(1)} \), I obtain a new vector of probabilities by applying a single iteration in the best response mapping, that is, \( \hat{\theta}_1 = \Psi_i\left( \hat{\theta}_{2s}^{(1)}, \hat{\theta}_0 \right) \). Then \( \hat{\theta}_{2s}^{(2)} \) maximizes the pseudo likelihood function \( \ln(\theta, \hat{\theta}_1) \) and so on. In the end, the NPL algorithm generates a sequence of estimators \( \hat{\theta}_{2s}^{(K)}: K \geq 1 \), where the \( K\)-stage estimator is defined as:
\[ \hat{\theta}_{2s}^{(K)} = \arg \max_{\theta \in \Theta} \ln(\theta, \hat{P}_{K-1}) \]

and the probabilities \( \{\hat{P}_K: K \geq 1\} \) are obtained recursively as:

\[ \hat{P}_K = \psi(\hat{\theta}_{2s}^{(K)}, \hat{P}_{K-1}) \]

4. Data
   (to be filled)

5. Results
   (to be filled)

6. Conclusion
   (to be filled)
References


