INDIAN INSTITUTE OF MANAGEMENT KOZHIKODE

Working Paper

IIMK/WPS/284/QM&OM/2018/28

May 2018

Does Story Really Matter In The Movie Industry? : Pre-Production Stage Predictive Models

Krishnan Jeesha[1]

Sumod S D [2]

Prashant Premkumar [3]

Shovan Chowdhury [4]

[1] Doctoral Student, Marketing, Indian Institute of Management Kozhikode, IIM Kozhikode PO, Kozhikode – 673570, E-mail: krishnanj08fpm@iimk.ac.in

[2] Doctoral Student, Strategy, Indian Institute of Management, Kozhikode, IIM Kozhikode PO, Kozhikode – 673570, E-mail: sumods08fpm@iimk.ac.in

[3] Doctoral Student, Quantitative Methods and Operations Management, Indian Institute of Management Kozhikode, IIM Kozhikode PO, Kozhikode – 673570, E-mail: nairp08fpm@iimk.ac.in

[4] Associate Professor, Quantitative Methods and Operations Management at the Indian Institute of Management Kozhikode, Kozhikode, India. IIMK Campus P.O., Kozhikode, Kerala 673570, India; Email: shovanc@iimk.ac.in, Phone Number (+91) 4952809119.

Abstract

The objective of the study is to develop a parsimonious model to predict the box office success of a Bollywood movie before its release based on historical data. A movie is considered successful if it is able to generate a ROI (return on investment) higher than the weighted average risk-free rate of return. The performance of a total of 447 movies over a 9 year period were examined. A set of variables that were identified as determinants of a movie's box office success by previous literature were tested for their applicability in the Indian context. In addition, certain variables that were unique to the Indian movie industry were investigated for their influence on the box office success of movies. The results demonstrate that factors like budget, screen count, genre, and release period all have significant influence on the outcome of a movie at the ticket window. However, contrary to popular belief, the historical box office performance of the lead actor, director or music director, and retelling of an existing narrative in the form of a remake were not found to add footfalls during the release of the movie.

Introduction

Movies are said to be all about 'selling dreams'. Yet, many a times the dreams of those associated with a movie gets shattered when it fails to find audience appreciation and moolah. This is because movie making is an art as well as a business. Like every art form, the appeal of a movie varies from viewer to viewer. Having said that, movie making is also a business involving huge investments and it is vital to ensure returns on these investments. However, unlike other enterprises, movie business sees the launch of new products every other week and relies on these new entrants to bring in majority of the turnover. The shorter shelf life and 'experiential' nature of the product (Eliashberg & Sawhney, 1994) compounds the complexity of problems faced by investors. Hence, several researchers have tried to discover the formula for financially successful films.

Studies to predict the box- office success of movies have primarily focused on Hollywood- the American motion picture industry which is the largest movie industry in the world and accounts for about $11.4 bn in domestic ticket sales and another $ 27.1 bn in global box office collections (MPAA, 2017). Some authors have also examined box- office predictors in other developed countries like UK and Australia (Elliott & Simmons, 2008; McKenzie & Walls, 2013). However, these markets are also heavily dominated by Hollywood movies. There have been very few studies that provide an off-Hollywood perspective on movie success factors (for exceptions see (Bagella & Becchetti, 1999)). India, on the other hand, provides a unique setting to test the models developed in the context of Hollywood.

The Indian film industry is bigger than Hollywood in terms of number of films produced with around 1,500 films produced every year. Being home to the second largest population in the world, it also enjoys one of the highest footfalls in the cinema halls. Furthermore, unlike Hollywood which generates more than 70% of its revenues from global sales, Indian movies generate close to three fourths of their revenue from the domestic box office. Nevertheless, Indian movie industry is split into several regional industries with movies being produced in more than 20 different languages. Of this, Hindi movies account for around 43% of the sales from roughly 150-200 number of movies (Deloitte, 2016) 50% coming from regional language movies and remaining from international (mainly Hollywood) movies. The disintegrated nature of the industry, and high regionalization and linguistic polarization makes it difficult to produce content appealing to a pan-Indian audience. The rampant piracy also reduces the window available for generating theatrical revenues and creates added pressure on the movie makers to recoup their investments at the earliest. Hence, analyses of the determinants of box office success of Hindi movies becomes important.

The rest of the paper is arranged as follows. The next section presents a survey of literature. This is followed by a description of the variables used in the study. Following section gives the results of the study. The paper ends with the discussion and conclusion sections.

Literature Review

Several studies have examined the determinants of box –office performance of movies. In one of the earliest studies, (Litman, 1983) explored the influence of genre, Motion Picture Association of America (MPAA) rating, budget, star value of the lead actors, distributors, release date and peer evaluation in the form of award nominations and awards on the box office performance of Hollywood movies and found that budget, science fiction genre, major distributor, release dates coinciding with the Christmas season, OSCAR award nominations and winnings were significant determinants of a movie's box office performance. Litman and Kohl (1989) replicated and expanded the initial work and demonstrated that factors like sequels, reviewer ratings and industry association ratings also influence the performance of movies. Contrary to the previous findings authors found that Summer releases had more chances of success unlike Christmas releases, and production costs, award nominations and winnings were less influential in the changed climate. Sochay (1994) repeated the study by considering all the 19 independent variables used by Litman along with a concentration ratio variable that depicted the competition faced by each film at the box office during its release. The author argued that contingent factors like competitive intensity would decide the best period to release a movie and explained that this would vary in different years due to distributor's predilection to avoid competition. He also introduced a new measure of movie's performance- the length of run and argued that it provides a more objective and reliable evaluation of performance as against the distributor rental figures that were used in Litman's studies.

Prag and Casavant (1994) used a different measure of performance- the box office figures to determine the success of a movie and showed that quality and marketing expenditures are vital determinants of a movie's success. They also revealed that film ratings, budget and star power matter more in the absence of marketing. Star power, awards, and budgets were found to be positive indicators of advertising spending. While these studies were predominantly from the perspective of the producers, Sawhney and Eliashberg (1996) developed a parsimonious model to predict the gross box office revenue of a motion picture to help the exhibitors in maximizing the yield from their exhibition capacity. Authors found that presence of major stars, familiarity of sequels and positive ratings from reviewers would positively influence the box- office potential of a movie. Another interesting result from the study was that prediction of the spread of revenue across the life cycle of a movie was almost impossible. The study also showed that the predictions on box office performance greatly improves with the availability of additional data from the initial few days of the movie's release which is very useful to the exhibitors to decide how long to play the movie in the theaters.

De Vany and Walls (1999) examined the probability of success of a movie based on stardom, budgets, genre, rating, opening screen counts, and year of release and survival time. A hit movie was defined as one with earnings in excess of $ 50 mn and the probability to achieve this figure was calculated. The independent variables were all found to be statistically significant. Moreover, they discovered that failure rate of movies was dependent upon time and long-runs were not a guarantee for success as the revenues tended to follow a convex path. Also, the positive impact of higher number of opening screens were found to peter out if the movie was rejected by the audience. Similarly, De Vany and Walls (2004) found that studios that follow 'blockbuster' strategies should focus on factors like budget, star power and number of opening screens to get better results. Walls (2004) highlighted the significance of 'crimedriven-action-adventure’ story lines to transcend the barriers of language and culture and bring bigger profits.

Ravid (1999) explored the role for stars and other potential information signals on the success of movies. Results indicated that stars have no influence on the profitability of movies. Author explained this using the ‘rent capture’ hypothesis which argues that stars claim remuneration equivalent to their market value and thus increase the cost of production thereby reducing the chances of better ROI even with higher revenues. Budget and critic’s ratings were found to significantly influence the performance of movies. However, higher budgets only helped to increase the revenue whilst bringing down the ROI of the movie. Similarly, sequels and MPAA ratings like G and PG also played a major role in determining the success of movies. In another study examining the risk related behaviour of executives in charge of large projects Ravid and Basuroy (2004) explained the proliferation of ‘R’ rated movies in Hollywood by demonstrating that these movies seldom lost money and their returns were more predictable even though their chances of earning mega bucks were limited.

Desai and Basuroy (2005) examined the interactive influences of genre, star power, and critics’ reviews on the market performance of movies. The results showed that as audience familiarity with the movie’s genre increases, star power and valence of critics’ reviews become less influential for its box office success. For less familiar genres, higher star value and positive critics’ reviews made a significant impact on movie’s performance. It was also found that the valence of critics’ reviews was more important for bigger star powered movies. In another study that investigated the ex-post-performance of movies at the box office, Deuchert, Adjamah, and Pauly (2005) revealed that Oscar nominations had a positive impact on the movie revenues and length of run at the box office whereas winning an award contributed very little in terms of extra revenues. Similarly, Chang and Ki (2005) focused on the ‘experience good’ property of movies to find the factors influencing the consumption of movies at the box office. The study revealed that Sequel, actor, budget, genre (drama), MPAA rating (PG and R), release periods (Summer and Easter), and number of first-week screens were significantly related to total box office performance.

Brewer, Kelley, and Jozefowicz (2009) examined the performance of Hollywood movies using domestic box-office revenue. They studied a wide variety of variables like budget, maximum number of screens in which the film was exhibited, economic factors like personal income and consumer price index for movie tickets, MPAA rating, genre, star power, sequels, critics’ rating, word-of-mouth, and award nominations that could influence the box office performance of movies. The study distinguished between information available to the public before and after the release of the movie. Significant positive determinants in the ex-ante regression model were budget, summer and holiday release dates, critical reviews, sequels and genres whereas budget, the peak number of screens, sequels, critical reviews, summer and holiday releases, word-of-mouth, award nominations and star power were found to have significant influence in the ex-post regression models.

Nelson and Glotfelty (2012) conducted a large sample study of the box performance of movies across 9 countries to evaluate how the star power of the talent associated with a particular movie influences its box office outcome. They operationalized star power as a continuous measure by counting the number of visits to the star's page in the IMDB (Internet Movie Database) website at the beginning of the year in which the movie was released. The study showed that star power of the lead actors had a major influence at the ticket windows and the synergistic effect of multiple stars would further improve the box office performance of movies. However, director's star power was not found to have any influence on the eventual outcomes the films. Another study by Bohnenkamp, Knapp, Hennig-Thurau, and Schauerte (2015) used a dataset of 2168 movies to investigate the box office potential of remakes using a sensations-familiarity framework of hedonic media. Authors argued that though remakes were not able to guarantee abnormal returns, they reduced the risk of failures. They also showed that remakes of movies with medium awareness, medium/low brand image, medium recency, movies belonging to horror genre and movies that were not inextricably linked to a specific artist or director had greater chance of success if they also had a slightly different treatment of their content from the parent movie.

As is evident, the research focus has been on the critical success factors of Hollywood movies. However, some researchers had also looked at the determinants of a motion picture's success in other contexts. For example, Bagella and Becchetti (1999) studied the box office results of Italian films over a 12 year period and demonstrated that the ex- ante popularity of actors coupled with that of directors were principal reasons behind the success of movies. They also showed that movies belonging to comedy genre had greater chances of success at the box office. Similarly, McKenzie and Walls (2013) studied the performance of Australian films at Australian box office and found that in spite of the higher levels of advertising and larger number of release theatres Australian films under-perform in terms of opening week and cumulative box-office collections. Elliott and Simmons (2008) used a data set of 527 movies to assess the relative importance of different quality signals (like budget, star power, highprofile directors, award nominations) given out through advertising on the success of films released in the United Kingdom. Films distributed by major US studios, those with 'U' ratings, sequels and award nominated films were found to have better performance. Similarly, critic's ratings, budgets and US box office collections (for films that had been previously released in US) were also found to influence the box office outcomes. Likewise, Fetscherin (2010) investigated the determinants of Indian movies released in US and UK and found that genre, movie rating, number of release screens influenced the performance of movies in these countries whereas star power, director power, distributor power and audience reviews were found to be of little significance.

An analysis of the literature shows that studies on the determinants of movie performance have focused extensively on the Hollywood motion picture industry. Those studies that happened outside the ambit of Hollywood were also limited to countries which were greatly influenced by the soft-power of US movie industry. Moreover, most of these studies did not differentiate between predictors of success before and after the release of movies. For example, though critics rating may decide the eventual box office performance of a movie, it is not helpful to a producer who is in the initial stages of conceptualizing a film project. The present study focuses on the predictors of movie success from the investor's perspective and hence has much more practical implications. To the best of our knowledge, this is one of the earliest academic studies that systematically analyses the determinants of a movie's success in the century old Bollywood industry. Additionally, unlike the extant literature which focuses on homogenous markets, it investigates the determinants in the domestic context of India with its extreme diversity of cultures, languages and sub-national and regional politicizations. Hence, the contributions from the study are expected to extend the boundaries of knowledge in this domain.

Methodology

The purpose of this study is to come up with a model that will predict based on the values the variables take, whether a movie will be successful (profitable to the producer) or a failure (a loss-making venture). Since we have data pertaining to the overall budget for a movie and the worldwide gross revenue figures, we could calculate the return on investment that a film might make instead of classifying it just as a hit or flop.

The Box Office classifies a movie based on its returns in to Super-hit, Hit, Plus, Average, Losing and Flop. However, rather than classifying in to multiple categories, we classify each movie as either a hit (return on investment at more than the risk-free rate of return) or a flop (otherwise). The risk-free rate of return for each year is available on Reserve Bank of India (RBI) website. Earnings of each movie (collections from India) will be adjusted for inflation to arrive at the classification. Inflation figures for each year is taken from RBI website.

We test both these approaches in this paper. We use multiple regression to test the former, while the latter approach is tested using logistic regression. The variables considered for both these approaches are listed below. All the independent variables remain the same for both the approaches. The dependent variable in case of the multiple regression is the ROI whereas in the case of logistic regression is the classification of the movie in to two categories – Hit or Flop.

Variables

Dependent Variable

ROI (For Multiple Regression): ROI is calculated as worldwide profits (Worldwide Gross Revenue – Budget) as a percentage of the Budget for the movie.

Classification (For Logistic Regression): Various movie rating agencies classify movies in to four or five categories based on the collection over a specific period of time. For example, according to the classification available on www.koimoi.com, movies are categorized into Super-hit, Hit, Average, Plus, Losing and Flop. The basis of classification is as follows:

Super-Hit - Film which has at least 150% returns on the investment

Hit - Film which has between 100% and 150% returns on its investment

Plus - Film which recovers investment and yields some profit up to a maximum of 100% returns

Average - Film which just recovers the investment

Losing - Film which does not recover the investment but loses less than 50% of it

Flop - Film which loses 50% or more of investment.

However, for the purpose of Logistic Regression, we classify the movie in to Hit or Flop based on whether the ROI is positive or negative respectively.

Independent Variables

Bankability of Lead Star: The success or failure of a movie depends to a large extent on the star cast of the movie. Particularly in the context of Bollywood movies, where a majority of the movie storylines revolve around the lead actor's character, the chances of a movie to succeed depends largely on who plays the lead character. An actor who has consistently delivered successful results in the recent past is more likely to attract more viewers as compared to someone who has delivered flops recently or someone who is a new entrant to the industry. Accordingly, we define a variable, bankability of an actor, which is measured as the moving average of the ROI of movies acted by the lead star in previous three years.

Previous Film of Lead Star: Similar to the recent track record of the lead actor, the success or failure of the actor's most recent movie also has an impact on the number of viewers hitting the theatres for watching the actor's next movie. This variable is defined to represent the commercial status of the actor's previous film. We take the ROI of the lead star's previous movie to represent this variable. However, in the case of debutants, the previous film is not applicable and hence we take an ROI of zero.

Hit-Director: Apart from the presence of a bankable lead actor, movie-goers also follow the track record of various directors. Consequently, a bankable director is more likely to maintain his success level not only due to the likelihood of him maintaining his quality of work, but also the viewers expectations from the director leading them to watch the movie. This variable too is measured as the moving average of the ROI of movies directed in previous three years.

Director's Previous Film: Just like the lead actor's recent performance has an effect on the increase in number of viewers of subsequent movie, the Director's most recent movie's result also has an impact on the number of viewers that will watch the movie. This variable is used to represent the commercial status of the director's previous film. We take the ROI of the director's previous movie to represent this variable. However, in the case of debutants, the previous film is not applicable and hence we take an ROI of zero.

Genre: Genre represents the style or category of art. Genre is an important factor which decides the reach of a movie and its eventual box office success. For example, it is believed that a family drama has a wider reach as it attracts all segments of audience when compared to an action movie which is preferred by youth and male audience. Horror movies often end up with

Adults Only certification which limits the audience it can attract. The data has been collected from Boxofficeindia.com

Music: Music is an integral part of Hindi movies. A popular musical score greatly aids the success of a movie. Hence, we consider this as an important factor to predict the success of movies. This is captured based on the historical performance of the Music Director. We take the ROI of the music director's previous movie to represent this variable. However, in the case of debutants, the previous film is not applicable and hence we take an ROI of zero.

Release Period: Release period represents the occasion during which the movie is released. Two types are identified- Festival and Non- Festival releases. Festival releases refer to those movies that are released during the occasion of major religious festivals like Diwali, Dussehra, Id, Christmas, Holi etc. or during national holidays like Independence Day, Gandhi Jayanti, Republic Day, New Year etc. Since movie going is a fun activity, the likelihood of this happening is more around a festival/public holiday. Moreover, these movies get extended weekends or holidays during the first week of release when maximum business for the movie is generated.

Franchise Movies: These movies are released as sequels/prequels of previous movies or take some reference (spin-offs) from previous movies. A prequel/sequel/spin-off provides better familiarity and recall value for the audience which improves the probability of success. The information can be collected from a content analysis of the news reports and reviews of the respective movies.

Remake: These movies have stories adapted from regional Indian movies or Hollywood movies or old Bollywood movies. Often movies that are successful in one language/region are considered for remake as it is believed that they would find patronage with a new set of audience and hence have a greater chance of success at the ticket counter. The information can be collected from a content analysis of the news reports and reviews of the respective movies.

No. of screens: The no. of screens in which the movie is released. This has a direct impact on the revenue from Box Office as higher number of screens lead to higher revenues and lesser payback period for the investment in the movie. Moreover, expensive movies need a higher catchment area to increase their chances of success. Since there is a wide variation in this number between movies and over the years (due to the increase in the screen availability on account of the spurt of multiplexes and the technology of digital projection supported by firms like UFO movies) we normalize the values by dividing the no. of screens in which a movie is released by the maximum no. of screens in which a movie was released in the particular year. The data is obtained from Boxofficeindia.com

Budget: The budget represents the production cost of the film which is available at Boxofficeindia.com

Results & Discussion

We ran a multiple regression model using the software package R. Since, we have data pertaining to the ROI of each movie, we proceeded with the multiple linear regression in an attempt to come up with a predictive model which would then allow a prospective movie producer to fit in the values of the variables in the model and check how much ROI can be generated from the movie that he/ she is planning to produce. The regression is run on a data set of top 50 grossing movies every year during the period from 2008 to 2016. However, of the 450 data points, three of the movies had missing data values and hence were removed, leaving 447 data points for the final analysis. The results of the multiple linear regression are given below:

```
Call:
lm(formula = ROI.Movie ~ Screens + Budget + Previous.film.of.star +
    Genre + Budget * Bankability.of.lead.star + Director.s.previous.film *
    Previous.film.of.star + Bankability.of.lead.star * Whether.Franchise +
    Budget * Release.period, data = Movie)

Residuals:
    Min      1Q  Median      3Q     Max
-2.0728 -0.7003 -0.1659  0.4889  4.6764
Coefficients:
```

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	-5.11E-01	2.58E-01	-1.979	0.04845	*
Screens	7.48E-04	1.23E-04	6.094	2.47E-09	***
Budget	-1.39E-09	4.95E-10	-2.807	0.00524	**
Previous.film.of.star	1.57E-01	5.22E-02	3.003	0.00283	**
GenreAdult	1.27E+00	1.05E+00	1.205	0.22885	
GenreAnimation	-5.56E-02	7.63E-01	-0.073	0.94202	
GenreComedy	2.42E-01	1.92E-01	1.256	0.20973	
GenreDrama	4.69E-01	1.87E-01	2.507	0.01255	*

GenreHorror	2.48E-01	2.96E-01	0.84	0.40159	
GenreLove Story	6.18E-01	2.38E-01	2.599	0.00967	**
GenreMasala	2.73E-01	3.43E-01	0.797	0.42618	
GenreRom - Com	5.50E-01	2.32E-01	2.374	0.01802	*
GenreSci Fi / Fantasy	-5.39E-01	4.24E-01	-1.27	0.20463	
GenreThriller	1.32E-01	1.97E-01	0.673	0.50141	
Bankability.of.lead.star	-1.33E-01	1.07E-01	-1.25	0.21188	
Director.s.previous.film	-7.56E-02	6.22E-02	-1.216	0.22451	
Whether.FranchiseYes	4.84E-01	2.22E-01	2.179	0.02989	*
Release.periodNormal	4.62E-01	1.76E-01	2.632	0.0088	**
Budget:Bankability.of.lead.star	5.13E-10	2.44E-10	2.102	0.0361	*
Previous.film.of.star:Director.s.previous.film	7.44E-02	3.19E-02	2.331	0.02022	*
Bankability.of.lead.star:Whether.FranchiseYes	-3.77E-01	1.84E-01	-2.053	0.04073	*
Budget:Release.periodNormal	-1.10E-09	3.75E-10	-2.938	0.00348	**

```
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.03 on 425 degrees of freedom
Multiple R-squared:  0.2456,   Adjusted R-squared:  0.2083
F-statistic: 6.587 on 21 and 425 DF,  p-value: 2.637e-16
                                                   GVIF Df GVIF^(1/(2*Df))
Screens                                        4.698986  1        2.167714
Budget                                         9.491423  1        3.080815
Previous.film.of.star                          1.632807  1        1.277813
Genre                                          1.871984 10        1.031847
Bankability.of.lead.star                       3.222717  1        1.795193
Director.s.previous.film                       2.031323  1        1.425245
Whether.Franchise                              2.204750  1        1.484840
Release.period                                 3.004073  1        1.733226
Budget:Bankability.of.lead.star                7.222696  1        2.687507
Previous.film.of.star:Director.s.previous.film 2.458972  1        1.568111
Bankability.of.lead.star:Whether.Franchise     2.157244  1        1.468756
Budget:Release.period                          3.217987  1        1.793875
```

The information in the table above also allows us to check for multicollinearity in our multiple linear regression model. All variables are within the tolerance level since VIF<10 for all variables.

From the above results, we can see that the Adjusted R-squared value of 0.2083 is low. However, given the high uncertainty in the movie industry, our model is able to explain almost 21% of the variation. The number of screens in which the movie is released is a significant variable contributing to the ROI of the movie and as expected, the ROI increases as the number of screens is increased. Contrary to our expectations, however, budget of the movie, despite being of significance, has a negative impact on the ROI. Increasing the budget of movie results in a decrease in the ROI. If at all the producer decides to go for a high budget movie, then it is better to go for an established and a bankable lead star. Also, the release date of the movie should be fixed during a holiday period so as to garner as much collections over the opening weekend as possible. Low budget movies on the other hand will be better off being released during a normal period rather than during holidays as there is less chances of competition from big budget star powered vehicles. Another interesting result we can observe is that producing a franchisee movie would be more profitable for a producer. However, surprisingly, a franchisee movie would be better off with a non-established star probably because the increased recall of the franchise helps to attract audience and which may be able to offset the lesser popularity of the star. Moreover, a non-established star may command lesser fee which reduces the investment in the movie and improves the ROI. ROI of the previous movie of the lead star has a positive impact on the ROI of the next movie the producer is planning to produce. The positive impact is further compounded by having a director who has had a successful previous movie. Genre-wise, Drama, Love-story and Rom-Com are significant variables which increase the ROI of the movie due to the higher acceptability among wider sections of audience. Contrary to our expectations though, comedy, masala and horror along with other genres turn out to be insignificant.

The above model can explain 21% of the variation. However, some producers might be just content with just knowing whether the movie will be profitable or a loss-making venture. Accordingly, we classified the movies in to these two categories. Movies with positive ROI over and above the weighted average risk-free rate of return are classified as profitable and those with corresponding negative figures as loss-making. Subsequently, we run a multiple logistic regression on the data set, the results of which are given below:

```
Call:
glm(formula = ROI.Movie ~ Screens + Budget + Previous.film.of.star +
    Budget * Bankability.of.lead.star + Director.s.previous.film *
    Previous.film.of.star + Bankability.of.lead.star * Whether.Franchise +
    Budget * Release.period + Genre, family = binomial, data = Movie)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.2657  -1.0319   0.4820   0.8857   1.6808
```

Coefficients:

	Estimate	Std. Error z	value	Pr(>|z|)	
(Intercept)	-3.09E+00	6.54E-01	-4.722	2.33E-06	***
Screens	1.72E-03	3.28E-04	5.234	1.66E-07	***
Budget	-5.59E-11	1.36E-09	-0.041	0.967143	
Previous.film.of.star	2.37E-01	1.26E-01	1.888	0.058984	.
Bankability.of.lead.star	2.08E-01	2.57E-01	0.808	0.419226	
Director.s.previous.film	5.16E-02	1.53E-01	0.338	0.735477	
Whether.FranchiseYes	5.60E-01	6.00E-01	0.934	0.350476	
Release.periodNormal	9.57E-01	4.06E-01	2.356	0.018451	*
GenreAdult	1.79E+01	2.40E+03	0.007	0.994043	
GenreAnimation	-1.40E+01	1.70E+03	-0.008	0.993406	
GenreComedy	1.08E+00	4.56E-01	2.379	0.01738	*
GenreDrama	1.51E+00	4.51E-01	3.344	0.000827	***
GenreHorror	1.77E+00	7.39E-01	2.388	0.016932	*
GenreLove Story	1.73E+00	5.91E-01	2.936	0.003327	**
GenreMasala	-2.89E-01	8.53E-01	-0.339	0.734456	
GenreRom - Com	1.81E+00	5.49E-01	3.3	0.000967	***
GenreSci Fi / Fantasy	-1.58E+01	8.00E+02	-0.02	0.984291	
GenreThriller	7.88E-01	4.61E-01	1.71	0.087321	.
Budget:Bankability.of.lead.star	-5.09E-10	6.68E-10	-0.761	0.446407	
Previous.film.of.star:Director.s.previous.film	-3.52E-03	8.95E-02	-0.039	0.968653	
Bankability.of.lead.star:Whether.FranchiseYes	-2.73E-02	5.29E-01	-0.052	0.958832	
Budget:Release.periodNormal	-2.42E-09	9.75E-10	-2.483	0.013013	*

```
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 581.91  on 446  degrees of freedom
Residual deviance: 475.90  on 425  degrees of freedom
AIC: 519.9

Number of Fisher Scoring iterations: 15
        ROI.Movie
glm.pred Loss Profit
```

```
  Loss      77     38
  Profit    82    250
Prediction accuracy
[1] 0.7315436
```

From the above results, we can see that like in the multiple linear regression, number of screens is a significant variable here as well and the chances of a movie being profitable increases as the number of screens increases. Budget, although, an insignificant variable in this model unlike the former case, for high budget movies, the likelihood of the success of a movie increases if it is released during holiday period. Keeping the budget constant, releasing a movie during the normal period increases success chances of the movie. The success of the previous movie of the lead star is a significant contributor to the success chances of the movie. The significance, however, is at 10% level, unlike the other significant variables which are at 5% level. Genre-wise, comedy, drama, horror, love-story and rom-com are all significant at 5% level and contribute positively to the success chances of a movie. Genre thriller is significant at 10% level though. The model has a prediction accuracy of 73% which is much better than taking a decision to produce a movie randomly.

Limitations

There was no single reliable database like IMDb and hence the data have been collected from multiple sources. Attempts were made to standardize data. Still there may be some effects on the results.

Many experts in the field opine that ‘Script’ and ‘Story’ are the most important determinants of a movie’s success. As the study tries to predict the success of the movie prior to its release these factors are not considered here. A comprehensive study which involves collecting feedback from the experts and audience would be required to understand this component.

Some hit movies got missed out in the analysis as the sample was limited to the movies of actors who have acted in at least three movies during the period. Also, actresses, their bankability and pull at the ticket counters was not considered in the study due to time constraints. Hence heroine-oriented movies which became box office successes got omitted in the study. Once included, more films would come into the frame which would make the findings more robust.

The study considers the box office collections against the amount spend on production and publicity of films. A major component of the film industry are the distributors and the exhibitors. Their profits can be analyzed only from the distributor's share and the per screen revenue of the movie respectively. Since these figures were not available, the analysis is not complete. This may create a situation where the film is profitable for the producer but may end up as loss making proposition for the other parties involved.

Most of the Indian movies fall in multiple genres. Hence the genre classification as used in the Hollywood movies is not applicable here. However, attempts have been made to identify the predominant genre of each movie and use that in the analysis. A new genre classification could be created for Indian film industry.

The popularity of music is derived based on the no. of YouTube hits for the songs. As the data was collected post the release of the movies, it may include hits that have happened after the time of release of films. Here the researchers make an assumption that no song had become popular after the movie's release and hence the variable may be construed as a good measure of popular music.

One of the determinants of movie's success is where the movie gets released. Two types of movies can be identified based on this. Movies released across the nook and corner of the country called the pan India releases or mass movies and those released only in the multiplexes of select cities and towns called the limited releases or multiplex releases. These niche movies catering to the tastes of the multiplex audience have greater chances of success if produced with reasonable budgets due to the targeted product, higher ticket rates and greater no. of shows in the multiplexes when compared to the mass releases. The variable 'extent of theatrical release' is not able to capture the percentage distribution of movies across these categories.

Only Indian Box office collections are considered for analysis. But several films, especially the lower budget films may be treated as profitable ventures for the producer if we consider the revenue from the sale of satellite (TV) rights, music rights and the collection from overseas markets. This is missed out in the analysis and if addressed could give a better sense of the Indian movie business.

There are a couple of interesting variables which have not been included in the current model but can certainly be explored in further studies

1. Target Audience: This variable indicates the target market for a particular movie. There are movies that are made for a pan-India audience from cities and rural areas. For example, a movie like Bajrangi Bhaijan has a story which is appealing to people across the country whereas a movie like Befikre has an urbane theme which would be more appealing to the metro audience. This decides the collections and success of the movie as niche movies would have only a small market to recoup their investment which needs to be factored in while budgeting for the production. As the data for this variable is difficult to obtain from the existing databases we are not attempting to test this in the present model.
2. Star Pairing: A lot of movie buffs rush to the movie halls to watch their favourite pair in a movie. Accordingly, this variable represents the bankability of the star pair based on their previous track record. This determines the buzz around the movie before its release and should positively influence the revenues of the movie.

References

Bagella, M., & Becchetti, L. (1999). The determinants of motion picture box office performance: Evidence from movies produced in Italy. *Journal of Cultural Economics*, *23*(4), 237–256.

Bohnenkamp, B., Knapp, A.-K., Hennig-Thurau, T., & Schauerte, R. (2015). When does it make sense to do it again? An empirical investigation of contingency factors of movie remakes. *Journal of Cultural Economics*, *39*(1), 15–41.

Brewer, S. M., Kelley, J. M., & Jozefowicz, J. J. (2009). A blueprint for success in the US film industry. *Applied Economics*, *41*(5), 589–606.

Chang, B.-H., & Ki, E.-J. (2005). Devising a practical model for predicting theatrical movie success: Focusing on the experience good property. *Journal of Media Economics*, *18*(4), 247–269.

De Vany, A., & Walls, W. D. (1999). Uncertainty in the movie industry: Does star power reduce the terror of the box office? *Journal of Cultural Economics*, *23*(4), 285–318.

De Vany, A., & Walls, W. D. (2004). Big budgets, big openings, and legs: Analysis of the blockbuster strategy. *Asian Economic Review*, *47*(2), 308–328.

Deloitte. (2016). *Indywood: The Indian Film Industry*.

Desai, K. K., & Basuroy, S. (2005). Interactive influence of genre familiarity, star power, and critics' reviews in the cultural goods industry: The case of motion pictures. *Psychology & Marketing*, *22*(3), 203–223.

Deuchert, E., Adjamah, K., & Pauly, F. (2005). For Oscar glory or Oscar money? *Journal of Cultural Economics*, *29*(3), 159–176.

Eliashberg, J., & Sawhney, M. S. (1994). Modeling goes to Hollywood: Predicting individual differences in movie enjoyment. *Management Science*, *40*(9), 1151–1173.

Elliott, C., & Simmons, R. (2008). Determinants of UK box office success: The impact of quality signals. *Review of Industrial Organization*, *33*(2), 93.

Fetscherin, M. (2010). The main determinants of Bollywood movie box office sales. *Journal of Global Marketing*, *23*(5), 461–476.

Litman, B. R. (1983). Predicting success of theatrical movies: An empirical study. *The Journal of Popular Culture*, *16*(4), 159–175.

Litman, B. R., & Kohl, L. S. (1989). Predicting financial success of motion pictures: The'80s experience. *Journal of Media Economics*, *2*(2), 35–50.

McKenzie, J., & Walls, W. D. (2013). Australian films at the Australian box office: performance, distribution, and subsidies. *Journal of Cultural Economics*, *37*(2), 247–269.

MPAA. (2017). Theatrical Market Statistics. Retrieved from https://www.mpaa.org/wp-content/uploads/2017/03/MPAA-Theatrical-Market-Statistics-2016_Final-1.pdf

Nelson, R. A., & Glotfelty, R. (2012). Movie stars and box office revenues: an empirical analysis. *Journal of Cultural Economics, 36*(2), 141–166.

Prag, J., & Casavant, J. (1994). An empirical study of the determinants of revenues and marketing expenditures in the motion picture industry. *Journal of Cultural Economics*, *18*(3), 217–235.

Ravid, S. A. (1999). Information, blockbusters, and stars: A study of the film industry. *The Journal of Business*, *72*(4), 463–492.

Ravid, S. A., & Basuroy, S. (2004). Managerial objectives, the R-rating puzzle, and the production of violent films. *The Journal of Business*, *77*(S2), S155--S192.

Sawhney, M. S., & Eliashberg, J. (1996). A parsimonious model for forecasting gross box-office revenues of motion pictures. *Marketing Science*, *15*(2), 113–131.

Sochay, S. (1994). Predicting the performance of motion pictures. *Journal of Media Economics*, *7*(4), 1–20.

Walls, W. D. (2004). Revenues, Profitability, and Returns: Clinical analysis of the market for mobster films. *International Journal of Business and Economics*, *3*(2), 93.

Research Office

Indian Institute Of Management Kozhikode

IIMK Campus P. O.,

Kozhikode, Kerala, India,

Pin - 673 570

Phone: +91-495-2809238

Email: research@iimk.ac.in

Web: https://iimk.ac.in/faculty/publicationmenu.php