Exploring the Influencing Factors of IP Film Rating by Sentiment Analysis and GMM

Yingxue Ma, *Mingxin Gan** *and Jiao Xv*

School of Economics and Management
University of Science and Technology Beijing, Beijing 100083, China
ganmx@ustb.edu.cn

Abstract

Recently, intellectual property (IP) film has become an important accessory for entertainment, and its rating has become the focus of quality evaluation. However, existing research seldom conducts study on influencing factors of rating. In this paper, we use sentiment analysis and generalized method of moments (GMM) to explore the factors that affect IP film rating. We take advantage of production, broadcast, genre and audience feedback to construct six explanatory variables, including actor influence, screenwriter participation, broadcast time, broadcast platform, genre, and adaptation satisfaction. We use LLC, IPS and Sargan tests to conduct variable stability test and model setting test. From the regression results of 134 IP films that obtained by sample filtering, the impact of each influencing factor on the rating is obtained. We found that short-term historical rating, actor influence, adaptation satisfaction and screenwriter participation positively affect current rating. While, long-term historical rating has a negative impact on current rating. In addition, broadcast time and broadcast platform have imposed positive impact on IP film rating, and genre has only a weak impact on rating. Our work provides advice for IP film producers, prompting them to improve quality by emphasizing celebrity effects and author participation.

Keywords: Film rating, GMM, sentiment analysis, influencing factor

1. Introduction

In recent years, intellectual property (IP) film has become an emerging product of the integration of literature and television industry. IP film is a kind of TV series recomposed from network novel, animation, game, drama or other works that have fans basis [1-2]. IP film is different from general TV series, since it has its own popularity due to the fact that there is already a market for original IP work. The original IP work regenerates fund worth by adaptation. Since the release of "Empresses in the Palace" on Netflix in 2015, IP film has attracted the attention of film and television producer. At the same time, the success of IP film made progress in game and entertainment industries, creating peripheral products, such as namesake mobile games and offline games, forming an industry chain. However, IP film has emerged homogeneous and low quality in its continuous development [3]. As an important measure for quality evaluation, IP film rating has attracted the attention of television producers and audience. Therefore, many producers pay for the ratings and clicks to avoid risks, which increases the chaos in IP film market. Film and television works have an important impact on audience's daily life and their psychology, guiding audience to form positive view of life, sense of worth and moral principle. However, the rating is audience-generated and cannot be controlled by humans. In order to achieve high rating and to better evaluate the quality of IP films, it is necessary to explore the factors that influence ratings, to achieve high rating by adjusting the relevant factors. Therefore, exploring the influencing factors of IP film rating can help producers improve the quality of TV products and avoid resource waste, promoting TV products to meet audience needs [4].

The representative research of influencing factors is the box office prediction of 687 North American films from 1981 to 1986, conducted by Litman et al. [5]. They analyzed the relationship between movie genre, movie rating and box office. Hennig et al. [6] found that there is a negative correlation between actors and box office of movies. Actors have an impact on both short-term and long-term box office. By contrast, the research on TV series rating is still insufficient. Rust et al. [7] investigated 5434 viewers during the prime time on Monday and Thursday. They pay attention to the association between the ratings of different TV channels and their TV programs, to predict the ratings in the same time on Wednesday and Friday. On this basis, Shachar et al. [8] used the logit model to estimate and test the viewing behavior of TV viewers.

URI: https://hdl.handle.net/10125/79677
978-0-9981331-5-7
(CC BY-NC-ND 4.0)

Generally, the research care about influence factors of TV series' ratings mostly take genre, broadcast time and broadcast platform as explanatory variables. Compared with research on movie box office that focus on actor influence, director and movie investment, it lacks the consideration of production and evaluation factors. Therefore, we explore the influencing factors of IP film rating. Since China's CCTV and Provincial TV are the main broadcast platforms of IP film in China, we selected the data from these two TV stations. We define six factors including actor influence, screenwriter participation, broadcast time, broadcast platform, genre, and adaptation satisfaction, from the aspect of production influence, broadcast influence, genre, and audience feedback. In order to obtain actual audience feedback, we collect user reviews on IP films from Douban. First, we calculate the sentiment score of user review, and normalize it as adaptation satisfaction. Second, we use the generalized method of moments (GMM) to analyze the impact of influencing factors on IP film rating. Finally, we give an importance evaluation of influencing factors and suggestions to improve the quality of IP film.

2. Variable selection and research methods

2.1. Variable selection and description

2.1.1. Predicted variable. We selected 134 IP films for research, which were first broadcast on or after January 1, 2011 and ended before December 31, 2018. The predicted variable (rating) is calculated by the weighted average rating of the same IP film broadcasted on different platforms, using ratings ranking, as shown in Eq. (1).

$$rating = \sum_{i=1}^{C} c_i \cdot \frac{C}{rank_i} \cdot raw\_rating_i \qquad (1)$$

Where, C represents the number of platforms. c_i is 1 if the IP film is broadcasted on the i-th platforms, otherwise c_i is 0.

2.1.2. Explanatory variables. The explanatory variables in this paper are listed in Table 1.

1). Actor influence and screenwriter participation. Taking the fact that audience will watch the film performed by their preferred actors into consideration, we take actor influence as one of the influencing factors in terms of production. Moreover, the participation of original author in script writing can help avoid the distortion of IP film, which is one of the advantages of TV film production. Therefore, we use the screenwriter participation as another influencing factor in terms of production.

2). Broadcast time and broadcast platform. Whether the broadcast time of an IP film is appropriate to audience's viewing time is an important factor that affects the rating. Therefore, we construct the variable of broadcast time, according to whether the broadcast time of the IP film belongs to prime time. At the same time, broadcast platform involves propaganda, investment and the scale of audience. We define broadcast platform as structural variable, taking Hunan TV as base group, constructing 9 dummy variables as broadcast platform factors.

3). Genre of IP film. Genre can distinguish different films and have a significant impact on their ratings. Genre used in our study is referred from the source website of IP films. Generally speaking, a film may belong to multiple genres at the same time. For example, the genres of Titanic include disaster, plot and love. We choose the first genre of each IP film as its genre. Therefore, we divide IP films into 7 categories: suspense, war, era, martial arts, palace, urban, and fantasy. We define the genre as a structural variable. Taking the fantasy as the base group, we construct 6 dummy variables as the genre factors.

4). Adaptation satisfaction. Since IP film is adapted from network novels, there have been some audience before adaptation. The feedback from original audience indicates the quality of adaptation. Therefore, we define adaptation satisfaction to reflect audience satisfaction on the adaptation of IP film.

Table 1. Definition of variables

Influence factor	Variable name	Symbol	Meaning
Production	Actor influence	*actor*	Actor's Baidu index before the day the TV series was released.
	Screenwriter participation	*scripter*	Whether the original author participated in the script writing.
Broadcast	Broadcast time	*time*	Whether the TV series broadcast during prime time.
	Broadcast platform	*dfws*	1 for Dongfang Satellite TV, 0 for other channels.
		zjws	1 for Zhejiang Satellite TV, 0 for other channels.
		jsws	1 for Jiangsu Satellite TV, 0 for other channels.
		bjws	1 for Beijing Satellite TV, 0 for other channels.
		ahws	1 for Anhui Satellite TV, 0 for other channels.
		szws	1 for Shenzhen Satellite TV, 0 for other channels.
		sdws	1 for Shandong Satellite TV, 0 for other channels.
		gdws	1 for Guangdong Satellite TV, 0 for other channels.
		scws	1 for Sichuan Satellite TV, 0 for other channels.
Genre	Genre	*dsj*	1 for urban, 0 for other genres.

		gtj	1 for palace, 0 for other genres.
		wxj	1 for martial arts, 0 for other genres.
		ndj	1 for era, 0 for other genres.
		xyj	1 for suspense, 0 for other genres.
		zzj	1 for war, 0 for other genres.
		qhj	1 for fantasy, 0 for other genres.
Audience feedback	Adaptation satisfaction	*satisfy*	The sentiment score of adaptation-related reviews given by viewers who know the plot of the original work.

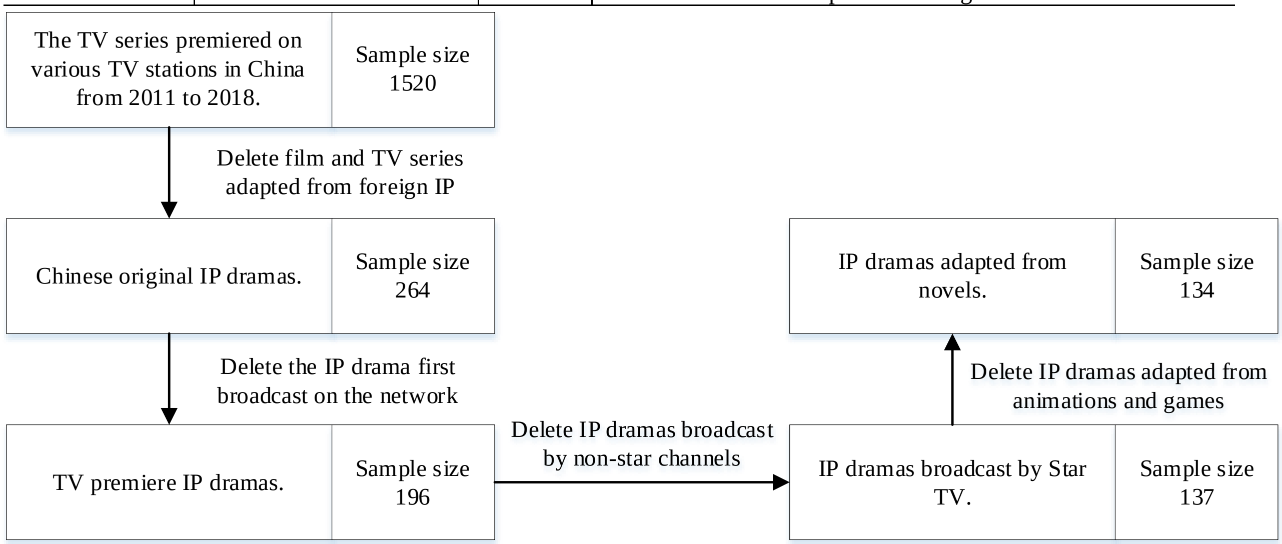

Figure 1. The process of sample filtering.

2.2. Data source

We crawled 1,520 TV series that broadcasted on China's CCTV and Provincial TV from 2011 to 2018. We determined a research sample that contains 134 IP films via artificial filtering, as shown in Figure 1. We used the unbalanced panel data of IP film rating for empirical analysis. In addition, we collected the raw data including editorial team, broadcast time, and user reviews from Baidu baike [1], Douban [2], and China Literature [3]. Actor influence is collected from Baidu Index [4]. The ratings are taken from CSM Media Research[5].

2.3. Research methods

2.3.1. Adaptation satisfaction calculation method based on sentiment analysis. The implementation of We Media platforms, such as video barrage, WeChat, and Weibo, enables audience to make real-time comments on films. Audience's comments in terms of adaptation indicate the adaptation quality and affect other's viewing decision. Therefore, we use the sentiment score of review to represent adaptation satisfaction. The sentiment dictionary based method is adopted to calculate the sentimental score of reviews [9-11]. The algorithm is shown in Table 2.

Table 2. The calculation steps of sentiment score.

Input:	Review texts $\boldsymbol{S}$
1:	Remove irrelevant contents, such as web links, pictures and emoticons in $\boldsymbol{S}$. Using *adaptation*, *original* and *comparison* as keywords to filter out user comments related to adaptation, obtain the pre-processed sentence $\boldsymbol{ST}$.
2:	Divide $\boldsymbol{ST}$ into short sentences and words, removing stop-words, etc. Obtain the pre-processed word set $\boldsymbol{S} = \{\boldsymbol{s_i}\}$.
3:	Calculate $\boldsymbol{s_i}$'s sentiment score $\boldsymbol{o_i}$.
4:	Calculate the weighted average of sentiment score $\boldsymbol{O}$ of $\boldsymbol{S}$. Obtain the sentiment scores of different IP films.

We map the sentiment score to [0, 1] interval, to obtain a normalized adaptation satisfaction. The larger the sentiment score, the higher the adaptation satisfaction is. In this paper, we use the china national knowledge infrastructure (CNKI) [6]semantic similarity method [12] to calculate adaptation satisfaction. First, we check whether the word before emotional word

[1] https://baike.baidu.com/
[2] https://www.douban.com/
[3] https://www.yuewen.com/#&about
[4] https://index.baidu.com/v2/index.html#/
[5] https://www.csm.com.cn/
[6] https://www.cnki.net/

belongs to modifier dictionary, and if so, we correct the sentiment score according to the intensity value of modifier in the dictionary. The set of modifier adverbs is defined as $w = \{<q_1, <at_1, at_2..>>, <q_2, <at_1, at_2..>>, <q_3, <at_1, at_2..>>\ldots\}$. Where, q represents emotional word, and a_t represents modifier adverb. The sentiment score $O(a_t)$ after adding sentiment adverb is calculated by Eq. (2). Examples of emotional word and modifier adverb are shown in Table 3.

$$O(at)=\sum_{i=1}^{n} q_i \cdot \prod_{j=1}^{m} at_j \qquad (2)$$

2.3.2. Econometric model. The research data comes from the rating of 134 IP films, n=134. In this paper, we established a panel data model using the weekly data [13-16], setting the sample size to 724, as shown in Eq. (3).

$$\begin{aligned} M: rating_{it} &= \alpha + \beta_0 satisfy + \beta_1 scripter + TYPE^{'}\gamma_k \\ &+ \beta_2 \log(actor) + CHANNEL^{'}\delta m + \beta_3 time \\ &+ \mu_i + \lambda_t + \varepsilon_{it} \end{aligned} \qquad (3)$$

Table 3. Dictionary of emotional word and modifier adverb

Word type	Category	Example	Intensity value
Emotional word	Strong commendatory	Perfect, obsessed, binge-watching…	+1.0~+1.5
	General commendatory	Good, okay, ok…	0~+1.0
	Strong derogatory	Vulgar, disgusting, offensive…	-1.5~-1.0
	General derogatory	Poor, single, fatigued…	-1.0~0
Modifier adverb	Extreme	Most, bottom…	+2.0
	Height	More, still, increasingly…	+1.75
	Moderate	Relatively, quite, comparatively…	+1.5
	Low	Slightly, briefly, somewhat…	+0.5
	Negative	Badly…	-1.0

Where, $rating_{it}$ represents the rating of the i-th IP film in the t-th week, and it reflects audience's preference. β is the parameter that will be estimated. *TYPE*' is a $1\times k$-dimensional structural variable vector that reflects the genre of IP film. γ is a $k\times 1$-dimensional parameter vector. *CHANNEL*' is a $1\times m$-dimensional structural variable vector that represents the broadcast platform. δ is a $m\times 1$-dimensional parameter vector. μ_i represents the unobserved individual effect. λ_t represents the unobserved time effect. ε_{it} represents the remaining random error.

We considered the dynamic characteristics of IP film rating, taking the impact of historical rating on current rating into consideration when constructing the model. The explanatory variable is determined by the lagged rating. Therefore, we established two dynamic panel data models, M_1 and M_2, as shown in Eq. (4) and Eq. (5).

$$\begin{aligned} M_1: rating_{it} &= \beta_0 rating_{i,t-1} + \alpha + \beta_1 satisfy + \beta_2 scripter + TYPE^{'}\gamma_k \\ &+ \beta_3 \log(actor) + CHANNEL^{'}\delta m + \beta_4 time \\ &+ \mu_i + \lambda_t + \varepsilon_{it} \end{aligned} \qquad (4)$$

$$\begin{aligned} M_2: rating_{it} &= \beta_{01} rating_{i,t-1} + \beta_{02} rating_{i,t-2} + \alpha + \beta_1 satisfy + \beta_2 scripter \\ &+ TYPE^{'}\gamma_k + \beta_3 \log(actor) + CHANNEL^{'}\delta m + \beta_4 time \\ &+ \mu_i + \lambda_t + \varepsilon_{it} \end{aligned} \qquad (5)$$

Where, $rating_{i,t-1}$ represents the rating of the i-th IP film lagged one period. $rating_{i,t-2}$ represents the rating of the i-th IP film lagged two periods.

3. Empirical analysis

3.1. Variable stability test

In order to alleviate the influence of spurious regression on empirical results, we conducted unit root test, LLC and IPS, to evaluate variable stability. The results show (see Table 4) that the first-order difference of all variables is stable.

Table 4. Results of variable stability test

Variable	LLC	IPS
rating	—	—
log(actor)	-11.8465***	-9.5747***
dfws	-10.7006***	-9.6049***
zjws	-8.9929***	-10.4706***
jsws	-5.4578***	-8.0399***
bjws	-8.1949***	-6.8502***
ahws	-3.5543***	**-2.2729*****
szws	-4.7514***	-10.8601***
sdws	**-1.8551*****	-3.2301***
gdws	-8.4409***	-18.7317***
scws	-3.9527***	-2.3582***
time	-8.1224***	-7.1247***
satisfy	-9.2774***	-10.2435***
dsj	-4.9931***	-2.8354***
gtj	-7.0829***	-4.3138***
wxj	-7.3174***	-4.4952***
ndj	-6.1339***	-3.6437***
xyj	-4.5835***	-2.9142***
zzj	-4.1352***	-2.4985***
scripter	-3.9645***	-2.3218***

*Note: *** means significant at the 1% level.*

3.2. Model setting test

We use generalized method of moments (GMM) to conduct model estimating, including differential GMM (DIF-GMM) and system GMM (SYS-GMM). We conduct Sargan test to verify the over-identification of instrumental variables in SYS-GMM. The results (shown in Table 5) show that the p-value of SYS-GMM (M_1) is 0.8156, and the p-value of SYS-GMM (M_2) is 0.9977. It accepts the hypothesis that "all instrumental variables are valid", indicating that there is no instrumental variable that related to the disturbance. The instrumental variables used in the model are reasonable. In addition, the p-value of AR (1) for the first-order autocorrelation test of error difference is 0.0000. It accepts the hypothesis that "the disturbance has autocorrelation", indicating that the disturbance difference in SYS-GMM has a first-order autocorrelation. The p-values of the second-order autocorrelation test AR (2) of the error difference

Table 5. Results of model setting test

Test aspects	SYS-GMM（M_1）	SYS-GMM（M_2）
Joint significance of TYPE	chi2(6) = 57.61 prob > chi2 = 0.0000***	chi2(6) = 103.26 prob > chi2 = 0.0000***
AR（1）	z = -6.0422 prob > z = 0.0000***	z = -5.5633 prob > z = 0.0000***
AR（2）	z = -5.0311 prob > z = 0.2368	z = -0.48357 prob > z = 0.6159
Validity of instrumental variables	chi2(38) = 62.2580 prob > chi2 = 0.8156	chi2(38) = 103.3915 prob > chi2 = 0.9977
Test number	724	724

*Note: *** means significant at the 1% level.*

are 0.2368 and 0.6159, respectively. It rejects the hypothesis that "the disturbance has autocorrelation", indicating that there is no second-order autocorrelation of disturbance difference in SYS-GMM. In a word, the results of the model setting test show that SYS-GMM is available to model dynamic rating. Therefore, we use two methods, DIF-GMM and SYS-GMM, for model estimation.

3.3. Results analysis of rating regression

The regression result is obtained by substituting each variable into M1 and M2. We find the following conclusions by investigating different influencing factors.

1). Historical rating has a significant impact on current rating. In the short term, historical rating has a positive impact on current rating. However, in the long-term, historical rating has a negative impact on current rating. In SYS-GMM (M_1), the p-value of first-order lagged $rating_{i,t-1}$ is 0.0000. It indicates that the rating lagged one period can significantly affect the current rating at a significant level of 1%. In SYS-GMM (M_2), the p-value of second-order lagged $rating_{i,t-2}$ is 0.0000. It shows that the rating lagged two periods is negatively correlated with the current rating at a significant level of 1%. In this paper, we define the period within a week as a short-term, and the period more than a week as a long-term. In SYS-GMM (M_1), the estimated coefficient of $rating_{i,t-1}$ is 0.3712. In SYS-GMM (M_2), the estimated coefficient of $rating_{i,t-1}$ is 0.3792. The estimations of these two coefficients are relatively close. It indicates that the average rating will increase by 3.712% or 3.792% when the average rating increases by 10% in the previous week. However, in SYS-GMM (M_2), the estimated coefficient of $rating_{i,t-2}$ is -0.1563. It shows that historical rating has a negative long-term impact on current rating. If the film's rating increases by 10% in the past two weeks, the average rating of this film will drop by 1.563% in this week.

2). Actor influence has a significant impact on rating. In terms of actor influence, the p-value of the regression results of SYS-GMM (M_1) and SYS-GMM (M_2) is 0.000. This means that the actor influence positively affects the rating.

3). Broadcast platform has a significant impact on rating. In the regression results of SYS-GMM (M_1) and SYS-GMM (M_2) on *CHANNEL*, the influence of 9 variables are significant at 1% significance level. Except for *zjws* and *jsws*, the estimated coefficients of the other variables are negative, which means that the ratings of Zhejiang TV and Jiangsu TV are all higher than those broadcast on the other 7 channels. It should be noted that the broadcast platform's influence on the rating is mainly related to its marketing capabilities. Therefore, the rating will be positively affected by the marketing capability of the platform.

4). Broadcast time has a significant impact on rating. In SYS-GMM (M_1), the estimated coefficient of time is 0.1743, while the estimated coefficient in SYS-GMM (M_2) is 0.2140. The results show that the rating of IP film broadcast during prime time (19:30-20:00) will be higher than those broadcast in other time (the average increment is approximately 0.1343 or 0.2140).

5). Adaptation satisfaction has a visible impact on rating. The estimated coefficient in SYS-GMM (M_1) is 0.3119. The estimated coefficient in SYS-GMM (M_2) is 0.3758. It shows that adaptation satisfaction has a greater positive impact on ratings than other variables.

6). Screenwriter participation has a visible impact on rating. In the regression results of SYS-GMM (M_1) and SYS-GMM (M_2) on *scripter*, the p-value is 0.000, which indicates that the participating of original author in screenwriting team have a positive impact on rating at a significant level of 1%.

7). The genre of IP film has a weak influence on rating. In DIF-GMM, the structural variable vector *TYPE* is directly eliminated. In SYS-GMM, the six variables contained in *TYPE* are not significant, but they passed the joint significance test. This shows that the genre of IP film has impact on their rating, but the impact is relatively small.

Table 6. Regression results of dynamic panel model

Variable	DIF-GMM（M_1）	DIF-GMM（M_2）	SYS-GMM（M_1）	SYS-GMM（M_2）
cons	0.2361*** (0.000)	0.4553*** (0.000)	12.3898 (0.439)	-13.9590 (0.312)
$rating_{i,t-1}$	0.1519*** (0.000)	0.0987*** (0.000)	0.3712*** (0.000)	0.3792*** (0.000)
$rating_{i,t-2}$		-0.6500*** (0.000)		-0.1563*** (0.000)
log(actor)	0.0148*** (0.034)	0.0213*** (0.000)	0.0076*** (0.264)	0.0217*** (0.001)
dfws	-0.2455*** (0.000)	-0.2459*** (0.000)	-0.2257*** (0.000)	-0.2197*** (0.000)
zjws	-0.0211*** (0.000)	-0.0247*** (0.000)	0.0226*** (0.000)	0.0314*** (0.000)
jsws	-0.0382*** (0.000)	-0.0329*** (0.000)	0.0315*** (0.013)	0.0362*** (0.000)
bjws	-0.0511*** (0.000)	-0.0507*** (0.000)	-0.064** (0.010)	-0.1136*** (0.000)
ahws	-0.1229*** (0.000)	-0.1357*** (0.000)	-0.1381* (0.080)	-0.1477* (0.095)
szws	-0.1583*** (0.000)	-0.1471*** (0.000)	-0.1822*** (0.000)	-0.2593*** (0.000)
sdws	-0.3850*** (0.000)	-0.3186*** (0.000)	-0.3357*** (0.003)	-0.3566*** (0.000)
gdws	-0.3992*** (0.000)	-0.3867*** (0.000)	-0.4920** (0.039)	-0.4796** (0.096)
scws	-0.5698*** (0.000)	-0.5927*** (0.000)	-0.5377*** (0.070)	-0.5977*** (0.013)
time	0.1224*** （0.000）	0.1247*** （0.000）	0.1743*** （0.000）	0.2140*** （0.000）
satisfy	0.2774*** （0.000）	0.2435*** （0.000）	0.3119*** （0.000）	0.3758*** （0.000）
dsj	—	—	-16.7818 (0.321)	13.8523 (0.363)
gtj	—	—	-11.5251 (0.229)	9.0967 (0.551)
wxj	—	—	-10.3884 (0.548)	23.0127 (0.142)
ndj	—	—	-5.2344 (0.294)	4.6301 (0.206)
xyj	—	—	-9.8856 (0.529)	11.8354 (0.356)
zzj	—	—	-13.1225 (0.294)	14.6301 (0.206)
scripter	0.0256*** (0.000)	0.0282*** (0.000)	0.0234*** (0.000)	0.0293*** (0.000)

*Note: 1. The number in brackets is the p-value of z-test of coefficient estimate. 2. "—" indicates that the variable is excluded from the regression results; 3. *, **, *** indicate 10%, 5%, and 1% significance level.*

4. Conclusion

In this paper, we construct six influencing factors, including actor influence, screenwriter participation, broadcast time, broadcast platform, genre and adaptation satisfaction, to evaluate the influence of production, broadcast, film genre and audience feedback on IP film rating. We use the sentiment dictionary based sentiment analysis method to calculate the adaptation satisfaction from audience's reviews. Both DIF-GMM and SYS-GMM are adopted to explore the influencing factors of IP film rating. Finally, we obtain the following conclusions: (1) historical rating, actor influence, broadcast platform and broadcast time significantly affect rating. (2) Adaptation satisfaction and screenwriter participation have a visible impact on rating. (3) The genre of IP film has a weak impact on rating. Our research calls on IP film producers to pay attention to the celebrity effect and the participation of original author in screenwriting work, to avoid the poor quality and homogeneity problem, helping to create specialized, differentiated and characteristic IP films.

As for future work, this work can be expanded by considering the sample of non-IP films to compare the difference between IP films and non-IP films. At the same time, the future work may consider more influencing factors such as actors and directors and its complex influence. Therefore, more advance methods, such as social network methods and deep learning models, may be adopted to capture the association between influencing factors and ratings.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Nos. 71871019, 71471016).

This work was funded by the Innovative Talent Support Foundation of University of Science and Technology Beijing.

References

[1]. Wang P. Research on the Interaction and Identification of the Interesting Groups in the IP Drama——Take the TV series "Sansheng III Shili Peach Blossom" as an example[J]. Home Drama, 2017(19):93-94.

[2]. Zhang H X. IP development and brand operation of domestic TV dramas in the "Internet +" era[J]. China Radio & TV Academic Journal, 2016(4):64-66.

[3]. Zhou X P, Han R N, Ling S. Research on the Influencing Factors of Internet Rating in the Context of Multi-screen Development——Take hit TV series as an example[J]. Chinese Journal of Journalism & Communication, 2014,36(12):114-129.

[4]. Zhang T, et al. Pre-broadcast ratings prediction method based on situational case-based reasoning[J]. Journal of Industrial Engineering and Engineering Management, 2020,34(06):156-164.

[5]. Litman B R. Decision - Making in the Film Industry: The Influence of the TV Market[J]. Journal of Communication, 2010, 32(3):33-52.

[6]. Hennig-Thurau T, Houston M B, Walsh G. Determinants of motion picture box office and profitability: an interrelationship approach[J]. Review of Managerial Science, 2007, 1(1):65-92.

[7]. Rust R T, Alpert M I. An Audience Flow Model of Television Viewing Choice[J]. Marketing Science, 1984, 3(2):113-124.

[8]. Shachar R , Emerson J W . Cast Demographics, Unobserved Segments, and Heterogeneous Switching Costs in a Television Viewing Choice Model[J]. Journal of Marketing Research, 2000, 37(2):173-186.

[9]. Chen L. An empirical study on the influencing factors of Chinese TV program ratings[J]. Industrial Economic Review, 2015, 4(4):152-160.

[10]. Yang X P, et al. Automatic construction and optimization of emotional dictionary based on Word2Vec[J]. Computer Science, 2017, 44(1):42-47.

[11]. Chen K Y, et al. Research on sentiment classification of hotel reviews based on sentiment dictionary[J]. Modern computer, 2017(6):3-6.

[12]. Hao Y Y, Ye Q, Li Y J. Research on the influencing factors of the usefulness of online reviews based on film review data[J]. Journal of Management Sciences in China, 2010, 13(8):78-88.

[13]. Li Q. GMM estimation of dynamic panel data model and its application[J]. Statistics & Decision, 2010, 2010(16):161-163.

[14]. Chen Jianbao, Cheng Suli. GMM Estimation of a Partially Linear Additive Spatial Error Model[J]. Mathematics,2021,9(6).

[15]. Zhang Hong Fan. Iterative GMM for partially linear single-index models with partly endogenous regressors[J]. Computational Statistics & Data Analysis,2021,156.

[16]. Kruiniger H. GMM Estimation and Inference in Dynamic Panel Data Models with Persistent Data[J]. Econometric Theory, 2009, 25(5):1348-1391.