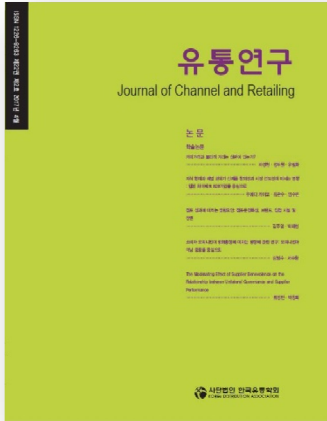


On: 30April 2017

Access Details: [Online: 2383-9694, Subscription:1226-9263]

Publisher Korea Distribution Association

Editorial office: School of Business, Konkuk University, Gwangjin-gu, Seoul, South Korea



## Journal of Channel and Retailing

Publication details, including instructions for authors  
and subscription information:

<http://www.kodia.or.kr/>

### “The Effects of Consumer Opinion on Box Office Performance in the Korean Movie Market: An Opinion Mining Application” \_학술논문

Bumsoo Kim<sup>a\*</sup>, Joo Hwan Seo<sup>b\*\*</sup>

a.Assistant Professor of Business Science, Sogang University

b.Assistant Professor of Marketing, DongA University

Online publication date: 30April2017

To cite this Article: Bumsoo Kim and Joo Hwan Seo (2017) ‘The Effects of Consumer Opinion on Box Office Performance in the Korean Movie Market: An Opinion Mining Application’, Journal of Channel and Retailing, 22(2): 65-91.

To link to this Article: [10.17657/jcr.2017.04.30.4](https://doi.org/10.17657/jcr.2017.04.30.4)

### PLEASE SCROLL DOWN FOR ARTICLE

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden. The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

## 소비자 오피니언이 영화흥행에 미치는 영향에 관한 연구: 오피니언 마이닝 응용을 중심으로\*

김범수\*\*, 서주환\*\*\*

본 연구는 2012년 1월부터 2014년 12월까지 국내에 개봉한 영화 중 총 537편을 선정하여 소비자의 오피니언이 영화흥행에 미치는 영향을 분석하였다. 이를 위해 영화매출에 영향을 주는 요인들을 크게 제작자와 소비자 측면으로 분류하였다. 첫째, 제작자 측면의 내재적 요인은 기존의 선행연구에서 진행된 스크린 수, 스타배우, 감독, 배급사, 시즌, 장르, 관람등급, 제작국가 등을 기반으로 하였으며, 선행 연구에서 살펴본 영화흥행 요인들의 영향을 확인하였다. 둘째, 소비자 측면의 요인분석은 온라인 구전(eWOM : electronic Word Of Mouth)을 통해 분석하였으며, 온라인 구전(eWOM)은 평점에 근거한 리뷰의 방향성, 리뷰의 수, 온라인 리뷰의 성향으로 측정하였고 이중 기존에 사용되지 않은 소비자 리뷰의 성향을 분석하기 위해 오피니언 마이닝 기법을 활용했다. 나아가 시점별 온라인 구전의 효과를 알아보기 위해 영화 상영전과 개봉후로 나누었다. 온라인 구전효과 중 개봉 전 평점은 영화매출에 유의한 영향을 주지 않았으나 개봉이후의 평점은 영화매출과 유의미한 관계를 보이는 것으로 나타났으며, 리뷰의 수 및 긍정적인 리뷰의 비율은 영화개봉 전과 이후 모두 영화매출에 긍정적인 영향을 주었다. 본 연구의 결과는 향후 영화산업의 의사결정자들이 영화의 유통환경을 이해하는데 일정 정도 기준을 제공할 것이며, 영화의 성공적인 매출성고를 위한 마케팅전략 수립의 근거가 될 것으로 보인다.

주제어 : 영화 매출, 오피니언 마이닝, 온라인 구전효과, 영화 리뷰, 소비자 성향분석

### I. 서론

21세기는 콘텐츠의 시대라고 할 만큼 콘텐츠산업의 매출과 수출이 꾸준히 성장하는 추세이다 (한국콘텐츠진흥원, 2015). 이러한 콘텐츠 산업은 ‘출판, 만화, 음악, 게임, 영화, 애니메이션, 방송, 광고, 캐릭터, 지식정보 및 콘텐츠솔루션’이라는 다양한 장르로 이루어져 있지만, 그 중에 특히 한류시장 확산 등 콘텐츠산업의 중요한 부분을 차지하는 것이 영화이다. 이에 본 연구는 다양한 콘텐츠 산업 중 영화산업에 대한 부분을 알아보고

자 한다.

영화산업은 통상 투자, 제작, 배급, 상영이라는 가치사슬을 형성하고 있으며, 여타 콘텐츠 산업에 비해 ‘고위험, 고수익’의 특성을 지니고 있다. <표 1>은 영화산업의 비용과 수입체계를 보여주며, 다른 콘텐츠에 비해 제작 후 비교적 단시간에 제품에 대한 성공 유무가 명확하게 판가름 나는 편이라 볼 수 있다.

그러므로 영화제작 관련업자들은 개봉 후 단기간 내에 이익을 발생시켜야 하며, 이에 따라 영화흥행에 영향을 미치는 요인들에 대한 관심도가

\* 본 논문은 동아대학교 교내연구비 지원에 의하여 연구 되었음.

\*\* 서강대학교 경영대학 경영과학 조교수 (bk4498@sogang.ac.kr), 주저자

\*\*\* 동아대학교 경영대학 경영학과 조교수(joohwans@dau.ac.kr), 교신저자

높아져 실제 많은 연구가 진행된 바 있다. 이런 다양한 선행연구들에 의하면, 영화 흥행을 예측하는 요인들로는 스크린 수, 배우 및 감독의 영향력, 배급사의 영향력, 제작국가, 평점, 장르, 댓글 수 등이 있다(Reddy et al. 1998; De Vany and Walls 1999; Elberse and Eliashberg 2003; 김은미 2003).

<표 1> 영화의 비용과 수익 구조

비용/수익	구분	정의
비용	순제작비	영화 제작에 실질적으로 소요되는 비용
	마케팅 비용 (Print & Advertisement)	배급에 필요한 프린트 비용 광고홍보비용
수익	극장매출	영화관 입장료
	부가수입	판권을 비롯한 입장료 외 수익

이에 본 연구는, 영화의 매출에 직접적으로 영향력을 미치는 여러 요소를 크게 제작자측면과 소비자 측면 요소로 나누어서 그 효과를 측정하는 한편 기존 연구에서 평균 평점으로 단순히 판단한 소비자 오피니언을 오피니언 마이닝을 통해 더욱 자세히 측정하여 그 효과를 분석하였다. 제작자측면은 기존의 선행연구에서 내재적 요인이라고 정의된 변수들; 감독, 유명배우, 스크린 수, 개봉시점, 관람등급, 배급사 및 장르 등을 변수로 사용하였다(Reddy et al. 1998; De Vany and Walls 1999; Elberse and Eliashberg 2003; 김은미 2003). 한편, 소비자측면의 요소로는 특히 정보원천과 관련해 구체적으로 온라인 구전(eWOM : electronic Word of Mouth)효과를 측정하였다. 온라인 구전은 크게 구전점수(Valence)와 구전리뷰 크기(Volume)로 나누어서 분석하였다. 왜냐하면 스마트 기기 및 인터넷의 비약적 발달로 인한 현 정보화 시대에 소비자들은 타인의 오피니언을 쉽

게 확인할 수 있을 뿐만 아니라 이를 의사결정에 반영하며, 이에 따라 이러한 온라인 구전효과와 관련된 연구의 필요성이 더욱 높아지고 있기 때문이다. 게다가 기존 연구와는 달리 온라인 구전효과를 분석함에 있어 평점이나 리뷰의 단순 규모만을 고려한 것이 아니라 처음으로 극성지수를 반영한 오피니언 마이닝 기법을 사용해 수많은 리뷰를 각각의 리뷰를 긍정적, 중립적, 부정적으로 분류하여 소비자 오피니언 효과를 더욱 세부적으로 측정하였다. 이를 위해, 본 연구는 2012년부터 1월부터 2014년 12월까지 국내에 개봉된 영화 537편을 대상으로, 총 59044개의 소비자 리뷰를 데이터 마이닝 기법으로 수집해 분석에 사용하였다.

이러한 연구는 향후 제작회사, 배급회사 및 상영회사와 같이 영화산업에 종사하는 의사결정자들에게 선제적인 정보를 제공할 수 있을 것으로 판단된다. 구체적으로 제작자 측면 요인의 분석은 영화의 제작단계에서 의사결정자들이 참고할 수 있는 정보를 제공하는 한편, 특히 소비자 오피니언, 즉, 긍정적인 평점과 리뷰의 분석은 영화를 홍보하는 시점에서 잠재 소비자들에게 영화의 긍정적인 정보를 제공하기 때문에, 이를 적극적으로 활용하는 마케팅 및 영화의 유통단계에서 사용할 필요가 있다. 이러한 다양한 정보를 바탕으로 영화산업 종사자는 영화 흥행을 위해 좀 더 효과적인 투자를 할 수 있으며, 또한 좀 더 효율적인 마케팅 전략을 세우는데 도움을 얻을 수 있을 것이다.

본 연구의 나머지 구성은 다음과 같다. 먼저 2장에서는 이론적 배경에 관한 선행연구 및 이를 바탕으로 한 연구가설을 제시할 것이다. 이어 3장에서는 본연구의 가설들을 검증할 구체적인 연구방법을 통해서 분석자료 및 변수들을 소개할 것이며, 4장에서는 실증분석을 통해서 실질적으로 제작자 요소와 소비자 온라인 구전 효과가 영화매

출에 얼마만큼 영향을 미쳤는지를 설명할 것이다. 마지막 5장에서는 연구의 결론과 시사점, 그리고 한계점을 통한 향후 연구방안을 제시할 것이다.

## II. 이론적 배경 및 연구가설

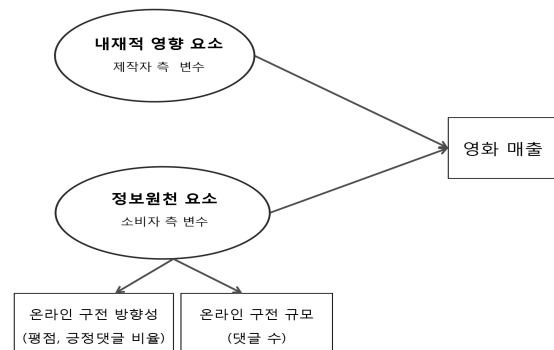
### 1. 영화 흥행에 미치는 내재적 요인들:

#### 제작자 요인

영화산업의 흥행요인에 대한 연구는 Garrison (1971)과 Simonet(1978)이 시작했으며 1980년대 Litman(1983)의 연구에서부터 관심을 받기 시작하였다. Garrison(1971)과 Simonet(1978)의 연구는 흥행 요인에 대한 표준적 지표를 제시했다는 점에서 의미가 있었으나, 표본이 작아 통계적 결과를 인정받기 충분하지 않았다. Litman(1983)의 연구는 회귀분석을 이용하여 영화 흥행에 영향을 미치는 요인을 분석했다는 점에서 방법론적 체계를 구축하였다. 나아가 Sochay(1994)는 영화의 흥행을 예측하는 것은 많은 변수가 존재하는 불확실성의 범주에 놓여 있기 때문에 측정이 매우 어렵다고 하였다.

이런 다양한 선행 연구들을 바탕으로 영화 흥행에 영향을 주는 요소를 크게 두 가지 요인으로 분류해 볼 수 있다. 이 두 가지 요인은, 크게 영화와 관련하여 제작자 측면에 해당되는 내재적 변수와 소비자 요인에서 접근하는 정보원천 변수로 나누어 볼 수 있다(Reddy et al. 1998). 첫째, 제작자 측면의 내재적 요인은 영화 고유의 속성과 관련된 기본적 변수이며, 여기에는 영화장르, 유명 배우, 유명 영화감독의 참여 유무, 관람객 등급, 스크린 수, 영화제작 국적 등이 있다. 둘째, 정보원천 요

인은 영화 외적인 부분에 해당되며 변수로, 광고, 평점, 영화 예고편, 리뷰 등이 있다. 이런 두 가지 요인들, 내재적 및 정보원천 변수는 영화 흥행에 필수적인 요소로 밝혀졌다(Litman and Kohl 1989; De Vany and Walls 1999; Elberse and Eliashberg 2003). 본 연구의 주목적은 소비자 오피니언이 영화 흥행에 미치는 영향을 살펴보는 것이나, 선행 연구에서 밝혀진 내재적 변수를 제외하고 분석을 진행할 경우 결과의 신뢰성에 대한 의문이 생길 여지가 충분하다. 이에 따라 아래의 <그림 1>과 같이 연구모형을 구성하였다.



<그림 1> 연구모형

내재적 변수에 대해 구체적으로 알아보면, 우선 주연배우 및 인기 감독의 영향력을 말할 수 있다.(Litman and Kohl 1989; Prag and Casavant 1994). Litman and Kohl(1989)은 인기 배우가 영화에 출연하면 인기 배우의 영향력 및 호감도가 높아져, 영화 흥행 가능성 또한 높아지게 된다고 주장했다. 하지만 이와 반대로, Prag and Casavant (1994)의 연구에서는 인기배우가 영화 흥행성과 별다른 영향을 미치지 않는다고 하였다. 감독의 영향력 또한 영화 흥행에 영향을 미칠 수 있다. 인기 감독이 만든 영화는 관객들의 호기심을 일으킬 수 있고, 소비자가 특정 감독에 대한 선호를 가지고 있거나 과거 특정 감독이 감독한 영화에

대해 긍정적인 경험이 있다면 영화 흥행에 긍정적인 영향을 미칠 것이다. 예를 들어, 안성아·김태준(2003)의 연구에서 감독의 영향력은 영화 개봉 초기 영화의 점유율에 긍정적인 영향을 미친다고 하였다. 그러나 감독의 영향력이 영화 흥행성과에 미치는 영향을 실증적으로 검증한 연구는 매우 적다.

그 밖에 배급사, 제작국가, 및 장르 등도 영화매출에 영향을 미치는 내재적 요인의 변수로 볼 수 있다 (Litman 1983; Litman and Kohl 1989; 박형현, 박찬수 2001; 유현석 2002; 이양환 외 2007). 우선, Litman and Kohl(1989)의 연구는 영화 개봉 직전 년도에 높은 수입을 올린 배급사일수록 극장주에 대한 협상력이 커지고 스크린 수를 확보하는데 유리해 진다고 하였다. 따라서 배급사의 영향력이 영화 흥행성과에 크게 도움이 되는 것으로 나타났다. 국내 연구를 살펴보면 이양환 외(2007)의 연구에서 영화 개봉 직전 년도에 높은 수입을 올린 영화를 많이 배급한 배급사일수록 영화 흥행성과에 영향을 미치는 것을 볼 수 있다.

영화의 제작국가 또한 영화 흥행에 영향을 미치는 요인으로 주목을 받아왔다. Litman and Kohl(1989)의 연구에서는 미국에서 제작한 영화가 다른 국가에서 제작한 경우보다 흥행에 긍정적인 영향을 미친다고 하였다. 하지만 박형현, 박찬수(2001)의 연구에서는 제작국가와 영화 흥행과의 관계는 없는 것으로 나타났다.

장르는 소비자들이 영화 선택과정에 있어서 중요한 요인이 중 하나로 영화에 대한 소비자의 선호를 반영할 수 있는 요인으로 주목받아왔다. Litman(1983)의 연구에서는 SF와 코미디 장르가 영화 흥행에 영향을 미치는 것으로 밝혀졌으며, 이양환 외(2007)의 연구에서는 코미디 장르가 영화의 매출에 영향을 미치는 것으로 나타났다.

이와 같이 영화흥행에 영향을 줄 수 있는 요소

는 무척 다양하며 각 요소의 영향력에 대한 분석 결과도 일관적이지 않은 모습을 보이니, 대부분의 선행연구들은 이러한 다양한 요인들이 동시에 복합적으로 영화흥행에 영향을 끼친다는 점에서 일치된 결론을 내리고 있다(Wallace et al. 1993; Sochay 1994; Litman and Kohl 1989). Wallace et al.(1993)은 장르, 감독, 주연 배우, 그리고 제작비 등의 변수가 영화 흥행에 영향을 주는 것으로 분석하였으며, Sochay(1994)의 연구에서는 주연 배우, 아카데미 수상실적, 개봉시즌, 스크린 규모, 코미디 장르가 흥행에 영향을 미치는 요인으로 분석했다.

국내에서도 영화의 흥행요인에 있어서 많은 연구들이 이루어졌으며 연구에 따라 다양한 시사점을 제시하고 있다(박형현, 박찬수 2001; 유현석 2002; 김은미 2003; 이양환 외 2007). 구체적으로 유현석(2002)은 1988년부터 1999년까지 개봉된 732편의 한국영화를 대상으로 흥행에 영향을 미치는 다양한 변수들을 분석하였는데, 연구 결과에 따르면 기존의 연구에서 확인된 내재적 변수들이 한국 영화시장에서도 효과가 있는 것으로 나타났다. 구체적으로, 주연 배우, 감독 및 제작자, 액션·코믹·사회풍자 장르, 15세미만 관람불가 등급의 영화가 상대적으로 흥행에 성공할 가능성이 높다고 나타났다. 김은미(2003)는 한국영화 흥행 결정요인에 관한 연구에서 다중회귀분석을 이용하였는데 이 연구에서는 출연배우의 수상경력 과 개봉 스크린 수, 배급사의 영향력이 영화 흥행에 영향을 미치는 유의한 변수로 분석되었다.

지금까지의 선행연구를 바탕으로 영화의 내재적 요인들은 영화를 제작하고 배급하는 제작자의 입장에서 영화흥행을 일으킬 수 있는 주요 요소들이라 할 수 있다. 이런 부분이 영화 매출의 주요 원인이 되며 영화흥행을 분석하는 연구모형에 기본적으로 포함되어야 할 필요성이 있다.

## 2. 영화 흥행에 미치는 온라인 구전효과 분석 : 소비자 요인

많은 사람들이 다른 사람들이 선호하는 제품을 선택하는 경향이 있으며, 이는 시장에서 다수에게 인기 있는 제품이 우수하다고 생각하기 때문이다.(Kramer et al. 2007; Salganik and Watts 2008, Khare et al. 2011). 구체적으로 Khare et al.(2011)의 연구에 따르면 사람들은 자신의 선호보다 자신이 속한 집단의 선호에 따라 제품을 선택한다고 분석하였다. 그리하여, 해당 집단의 선호가 소비자의 실질 구매에 영향을 미치며, 소수의 선호보다 다수의 선호가 더 큰 영향을 미친다고 했다.

이러한 구전 효과는 2000년 이후 스마트 기기 및 소셜미디어의 비약적 발전으로 인해 온라인 구전(eWOM)<sup>1)</sup>의 효과로 더욱 발전되었다. 최근 온라인 구전 효과에 대한 연구가 여러 분야에서 활발히 진행되고 있으며, 영화는 이러한 온라인 구전효과가 영향을 미치는 대표적인 분야 중 하나이다.(Liu 2006; Chintagunta et al. 2010; Zhao and Xie 2011; Fossen and Schweidel 2017). 정보의 비대칭성(Asymmetric Information)으로 인해 영화를 실제로 보기 전에 소비자들은 그 영화의 작품성이나 대중성을 판단하기 어렵다.

따라서 소비자들은 영화를 선택하기 전에 온라인 구전을 통해서 선택의 위험성을 줄이고자하는 성향을 보인다. 이와 관련된 선행 연구들을 보면, 온라인 구전은 영화흥행에 있어서 매우 중요하며 유의미한 변수인 것으로 분석되었다(Neelamegham and Chintagunta 1999; Dellarocas et al. 2007; Duan et al. 2008; Chintagunta et al. 2010). 최근에는 온라인 구전(eWOM)을 좀 더 구체적으

로 나누어서 영화흥행에 얼마나 영향을 미치는지에 대해 연구가 진행되고 있다(Godes and Mayzlin 2004; Chevalier and Mayzlin 2006; Zhao et al. 2013).

온라인 구전을 평가하는 지표는 온라인 평점 평균, 전체 리뷰어의 수, 리뷰어들의 분포 등과 같이 다양하나, 크게 소비자 오피니언을 보여주는 온라인 구전의 방향성(Valence)과 구전의 규모(Volume)를 측정하는 지표로 구분된다. 온라인 구전효과 중 방향성과 관련된 지표는 다양한 형태로 나타나지만 주로 5점 혹은 10점 척도를 사용한 평점이 주로 사용되고 있다. 이런 평점의 평균점수, 평점 분포의 차이 등을 통해 소비자들은 구전의 방향성과 다른 소비자들의 오피니언을 파악하고 이를 결정에 사용할 수 있는데 보다 구체적으로, Chintagunta et al. (2006)의 연구에서는 온라인 구전 효과 중 소비자들의 평균 평점으로 평가한 구전의 방향성만 영화 선택에 영향을 미친다고 했다.

반면 Liu(2006)는 온라인 리뷰의 성향분석을 통한 구전효과의 방향성과 영화매출 간 상호관계는 존재하지 않는다고 주장했다. 한편 배정호, 심범준, 김병도(2010)는 연립방정식을 이용한 영화 산업의 구전 효과 분석을 통해, 온라인 구전(eWOM) 점수가 영화 매출에 유의한 영향을 미칠 뿐더러 구전의 규모 또한 매출에 영향을 준다는 것을 보여주었다. 비록 측정 방식은 조금씩 다르나 온라인 구전 규모(volume)가 영화 흥행에 긍정적인 영향을 미친다는 것은 다른 연구(Ravid 1999; Liu 2006; Duan et al. 2008)에서 일관적으로 나타난 반면, 소비자 오피니언을 나타내는 구전의 방향성과 관련된 연구결과는 위에서 볼 수 있

1) 온라인 구전이란 인터넷 상에서 발생하는 커뮤니케이션을 말하며, 소비자들이 서비스에 대한 직,간접적 경험을 통해 얻은 긍정,부정적 정보를 교환하는 과정이다.(Sung et al., 2002). 온라인 구전 효과는 주로 온라인 구전 점수(Valance)와 온라인 구전 규모(Volume)로 측정하는데 연구자에 따라 결과가 상반된다.

듯이 매우 혼재되어 있다.

동일한 평가 프로세스 안에서 점수를 부여한다는 의미로 소비자가 주는 평점의 평균이라는 수치를 신뢰할 수 있다. 하지만 평점 분포가 발산하면 분산이 커지고 일부 소비자들의 극단적인 평점들이 평균 평점이라는 온라인 구전 방향성의 지표를 왜곡할 수 있다. 다시 말해서, 선행연구에서는 온라인 구전 요소의 방향성(Valence)을 측정하기 위해 짧은 시간에 손쉽게 수집할 수 있고 이미 수치화되어 있어 사용하기 편리한 평점을 선택하는 경향이 있으나, 소비자들마다 기준이 평가 기준이 매우 상이하고 극단적인 평점으로 인한 왜곡현상이 발생할 수 있다는 명백한 단점이 존재한다. 구체적으로 10점 척도에서 6점 이상의 점수를 긍정적인 시그널로 받아들이는 소비자가 존재하는 반면 8점 이상의 점수를 긍정적인 시그널로 생각하는 소비자들이 존재하는 등 개별 상황에서 평가 기준의 차이가 존재한다. 추가적으로 동일한 10점 척도 안에서 개인이 평점을 부여한다고 해도 모든 평가자들이 10점 척도를 온전히 활용하기 보다는 각 개인의 성향에 따라 10점 척도 중 일부만 사용하는 경우 또한 빈번하다. 예를 들어 평점의 분포가 6~10점에서 형성되는 개인이 있는 반면 1~10점을 모두 사용하는 평가자도 실존하는 등 개개인의 기준 차이로 인한 점수의 차이가 발생할 수 있다.

이러한 평점제도의 단점을 극복하기 위한 대안이 소비자가 작성한 리뷰의 내용을 분석하는 오피니언 마이닝<sup>2)</sup>이라 볼 수 있다. 글로 작성된 리뷰는 온라인 평점보다 자료수집과 분석에 더 많은 시간을 더 투자해야 하고 무수한 텍스트를 분

석하여 수치화 시켜야 한다는 단점이 존재하지만, 특히 표본의 크기가 크지 않을 경우 평점에 비해 일부 극단적인 평가의 영향을 덜 받는다는 점에서 비교적 객관적일 수 있고 소비자들의 비교적 정확한 오피니언을 측정하는데 적합하다고 (Chen and Zimbra 2010; Liu 2010) 볼 수 있다. 극단적인 예로 5명의 평가자중 4명의 평가자가 긍정의 의미로 7점을 부여하고 1명이 부정의 의미로 1점을 주었을 경우 평균 평점은 5.8이 되는 반면 긍정평가의 비율은 0.8이 된다.

이렇듯 평점이 긍정댓글 비율보다 극단적인 평가에 더 영향을 받으며 그러한 영향은 표본의 크기가 작을수록 더욱 심해진다. 일반적으로 온라인 리뷰 생태를 살펴보면 각 제품에 대한 리뷰의 규모가 크게 차이가 나며 일부 매우 성공적인 제품 외에 대부분의 제품에 대한 리뷰의 규모는 그리 크지 않은 것을 볼 수 있다. 이러한 모든 점을 고려할 때 단순히 평균 평점만으로 소비자의 오피니언을 측정하기 보다는 긍정평가의 비율도 같이 사용하여 조금 더 정확한 소비자오피니언을 측정하는 것이 중요하다고 판단할 수 있다. 이에 따라 본 연구에서는 온라인 구전효과의 방향성을 분석하는데 있어 소비자들의 평점에 추가적으로 기존 연구에서 활용되지 않은 텍스트마이닝을 통한 소비자의 오피니언 변수를 측정하여 분석에 사용한다.

덧붙여 기존 연구에서는 많이 다루지지 않았으나, 온라인 구전효과의 시점이 영화 개봉 이전 혹은 이후에 따라 영화흥행에 미치는 영향도 바뀌는 경우가 발생 할 수 있다. 영화 개봉 이전에는 짧은 영화 예고편, 광고 혹은 다른 매체를 통해 얻은

2) 오피니언 마이닝은 텍스트 마이닝의 한 분야로 최근 활발한 연구가 이루어지고 있다. 오피니언 마이닝은 소셜미디어나 콘텐츠 등으로부터 사용자의 의견을 분석하여 유용한 정보로 재가공하는 기술이다. 상품 및 선호도, 사회적 사건이나 이슈 등에 대한 대중들의 의견을 분석하고 사용자의 의견과 감정의 패턴을 이용하여 특정 정보에 대한 의견이 긍정인지 부정인지 중립인지 찾아내는 것이다. (Chen and Zimbra 2010; Liu 2010). 따라서 영화의 오피니언은 소비자의 감정을 분석하는 것이므로 오피니언 마이닝 분석이 적합하다.

정보와 같이 간접적 경험을 바탕으로 평점이나 리뷰를 남기게 된다. 반면 영화 개봉 이후에는 관람객들의 직접적인 경험을 바탕으로 소비자들이 해당 영화에 대한 보다 정확하고 구체적인 평가를 할 수 있다. 또한 일반적으로 온라인 구전의 규모를 살펴보았을 때 개봉 전 보다 개봉 후 더욱 활성화 되며, 온라인 구전이 크게 활성화되기 전인 개봉 전 평점의 경우 위에서 지적한 바와 같이 적은 표본수의 문제로 인해 지표의 신뢰성 문제가 실질적으로 나타날 수 있다. 이렇듯 작성 시점에 따라 온라인 구전의 본질적 특성이 다를 수 있다는 가정 아래, 본 연구는 영화 개봉 전과 개봉 이후의 온라인 구전이 영화 매출에 얼마만큼 유의미한 영향을 끼치는지 그리고 그 영향이 일관적인지도 알아보려고 한다. 위의 맥락을 종합하여 다음과 같은 가설 수립이 가능하다.

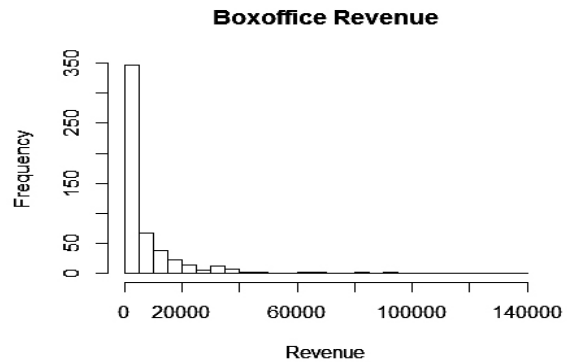
- H1: 온라인 구전효과(eWOM)는 영화 매출에 긍정적인 영향을 미친다.
- H2a: 온라인 평점(Valence)은 작성 시점에 상관없이 영화 매출에 긍정적인 영향을 미친다.
- H2b: 온라인 리뷰 규모(Volume)는 작성 시점에 상관없이 영화 매출에 긍정적인 영향을 미친다.
- H2c: 소비자 오피니언(Valence)은 작성 시점에 상관없이 영화 매출에 긍정적인 영향을 미친다.

### III. 연구방법

#### 1. 영화 흥행에 미치는 분석방법의 설계 및 변수

본 연구의 분석대상 영화는 2012년 1월부터 2014년 12월까지 개봉한 모든 영화 가운데 매출액이 30,000,000원 이상, 전국 스크린 수 60개 이상, 관객 수 5,000명 이상의 영화를 대상으로 한다. 이

러한 기준에 따라 영화진흥위원회 통합전산망(KOBIS, www.kobis.or.kr)에서 수집한 영화의 수는 총 537편이었다. 개봉 스크린 수가 60개가 되지 않거나 전국 관객 수가 5,000명 이하의 영화는 독립영화, 예술영화, 비상업영화 혹은 IPTV 시장으로 넘어가기 이전 형식적으로 개봉하는 영화가 대다수였다. 이런 영화들과 같은 경우 내재변수인 제작자 측 변수들은 존재하나, 네티즌 평점, 댓글의 수가 전무하거나 미미하여 정보원천 변수의 영향력을 분석하기가 어렵고 영화의 유통구조 또한 일반적인 영화들과 다른 점을 보아 소비자 오피니언의 효과를 파악하는데 적절치 못한 자료라고 판단하여 기존 연구(배정호 심범준 김병도 2010)와 같이 본 연구의 분석대상에서 제외하였다.



<그림2> 국내개봉영화 매출현황 (단위:백만원)

영화 흥행을 파악하기 위한 종속변수로 기존 연구에서는 영화 흥행 측정을 위해 매출액(Liu 2006; Chintagunta et al. 2010; Duan 2008) 혹은 실제 관객 수(유현석 2002)를 사용했다. 최근 국내에서도 고급 관람 좌석이 등장하는 부가가치 가격 정책을 하거나, 요일 혹은 시간별 다른 가격정책인 시간별 가격정책 등을 실시하고 있지만, 아직까지는 국내 영화 관람료가 전반적으로 일률적이라는 성향을 고려해 본 연구에서는 영화의 총 매출액을



선택했다.

수집된 영화들을 대상으로 매출액의 분포를 살펴보면, <그림 2>에서처럼 매우 소수의 영화만이 두드러진 성공을 거두는 반면 대부분의 영화는 낮은 매출을 보여 승자와 패자간의 격차가 큰 시장이라는 것을 알 수 있다. 더욱이 각 영화의 매출액은 평균 83억원 규모인데 반해 매출의 중간값은 21억원 규모로 격차가 매우 크고, 매출액의 왜도 또한 3.754 수준을 보이는 극단적 오른쪽 꼬리 분포의 모습을 보인다고 할 수 있다. 따라서 이러한 기초 통계분석에 따라 본 연구에서는 영화매출액을 로그변환(Chintagunta et al. 2010)하여 영화흥행을 측정하는 종속변수로 사용하였다.

### 1.1 내재적 변수의 정의 및 수집

영화흥행의 내재적 변수로 본 연구에서는 개봉 스크린 수, 감독과 주연배우의 영향력, 배급사 등급, 개봉시즌, 장르, 관람등급 그리고 영화의 제작국적을 선택하여 연구모형의 통제변수로 사용했다. 스크린 수와 같은 경우 소비자의 접근성에 영향을 미쳐 영화흥행에 직접적인 영향을 준다고 볼 수 있다. 일반적으로 영화상영의 프로세스를 살펴보면 초반에는 다수의 극장을 확보해 매일 상영을 하나 후반에는 초반 영화흥행 성적에 따라 확보 스크린 수의 변동이 심하고 특정 스크린에 다수의 영화를 교차 상영하는 등 분석의 왜곡이 발생할 수 있다. 이에 본 연구에서는 영화의 개봉 시기에 결정되는 개봉 스크린 수를 영화흥행의 내재적 변수로 사용했다.

영화에 참여하는 감독과 주연배우 개개인의 능력은 영화의 질에 직접적인 영향을 주기도 하지만 외부인들이 해당 영화를 바라보는 시각에도 간접적인 영향을 미쳐 흥행여부를 예측하고 분석함에 있어 주요변수로 사용되어 왔다.(Ravid

1999; Elberse and Eliashberg 2003; Ainslie et al. 2005) 기존 연구를 살펴보면 크게 제작자를 대상으로 한 설문조사(Ainslie et al. 2005), 수상경력과 누적관객 수(Ravid 1999) 등을 이용하여 스타감독/배우를 정의했다. 국내연구의 경우 제작 편수(김휴종 1998) 혹은 관객동원 수(유현석 2002)가 일정 기준을 초과한 배우를 스타배우로 분류한 후 터미변수로 정의하거나 지난 3년간 출연하거나 감독한 영화의 총 편수(김은미 2003)를 스타파워로 정의했다. 그러나 외국과 국내 영화의 수상경력을 객관적으로 동일선상에서 비교할 수 없고 영화계 종사자를 대상으로 한 설문이나 특정 기준을 부과해 감독/배우의 영향력을 측정하는 것은 매우 주관적일 수 있다는 현실을 고려하였다. 따라서 본 연구에서는 영화산업의 특성상 ‘성공적인’ 감독과 배우들에게 기회가 더 많이 주어진다는 점에 기안해 영화개봉 시점 이전 5년 동안 연출 혹은 출연한 영화의 수로 감독과 배우의 영향력을 평가하였다. 또한 공동주연이거나 영향력 있는 조연배우들도 고려해 주연배우 변수의 경우 영화의 크레딧에서 출연진에 기재된 첫 두 배우를 주연배우 1과 주연배우 2로 정의했다.

배급사의 영향력과 같은 경우 기존 연구(김은미 2003)에서는 일정 기간 동안 배급사가 배급한 영화의 편수로 측정했다. 그러나 많아야 1년에 1~2편 작품에 임하는 감독과 배우와 달리 배급사와 같은 경우 같은 기간 동안 다수의 영화를 배급할 수 있다. 뿐만 아니라 CJ엔터테인먼트 혹은 롯데엔터테인먼트와 같은 국내의 대형배급사의 존재로 개별 배급사간 배급하는 영화수의 격차가 커 단순히 배급 영화편수로 영향력을 측정할 경우 연구 결과의 오류를 초래할 수 있다. 이러한 문제점을 최소화하기 위해 본 연구에서는 단순 배급 편수보다는 배급사의 영향력을 등급화해 분석에 사용하였다. 보다 구체적으로 설명하면, 데이

터 수집 기간 동안 각 배급사가 배급한 영화 편수를 기준으로 서열을 책정한 다음 상위 20%에 속하는 배급사를 1등급 그 외를 2등급으로 분류한 후 더미변수를 사용하여 대형배급사의 영향력을 통제했다.

선행연구에서 영화흥행에 영향을 미칠 수 있는 다른 내재적 변수들로 언급된 개봉시즌, 장르 그리고 관람등급과 관련된 변수로는 일관된 기준이 존재하지 않고 영화 장르와 같은 경우 특히 복합 장르로 구분되는 경우가 많아 선행연구에서 정의한 변수를 선별적으로 모형에 추가했다. 김휴중(1998)은 한국영화 시장의 성수기를 여름시즌과 크리스마스/새해 겨울방학 시즌으로 선정하였고 이에 따라 본 연구에서도 개봉일 기준으로 12월~2월 그리고 6월~8월에 개봉한 영화를 성수기 시즌의 개봉영화로 정의하였다. 기존 연구(배정호 심범준 김병도 2010)에 따라 관람등급 관련 변수는 전체관람가/청소년관람불가/그 외 등급 영화로 장르 관련 변수는 액션(어드벤처)/코미디/그 외 장르 영화로 구분하여 분석에 사용하였다.

마지막으로, 기존 연구(배정호 심범준 김병도 2010)에서와 같이 영화 제작 국적에 따라 영화흥행에 미치는 영향을 측정하기 위해 한국영화와 외국영화를 구분해 변수로 사용했다.

## 1.2 정보원천 변수의 정의 및 수집

온라인 구전효과 분석은 다음과 같은 전처리 과정을 거친 후에 사용하였다. 일반적으로 영화의 개봉일자가 대략 결정이 되면 관련영화의 정보가 인터넷 포털사이트에 기재가 되고 이후 소비자들은 검색을 통해 관련영화의 정보를 탐색할 수 있다. 특정 영화에 대한 관심이 있는 소비자들은 영화에 대한 기대감/만족도를 댓글과 평점을 통해 표현할 수 있는데 이때 무분별한 댓글 혹은

평점을 제어하고 소수에 의한 여론 조작을 방지하기 위해 대부분의 포털사이트는 일차적으로 인증절차를 거쳐 해당 포털사이트에 로그인을 하도록 하고 있다. 본 연구에서는 “DAUM 영화”에서 각 영화의 사용자 평점, 댓글 수에 대한 자료를 수집하고 오피니언 마이닝을 이용한 댓글의 내용 분석을 통해 구전 방향성을 측정하는 추가적 소비자오피니언 변수로 사용한다. 여러 웹 사이트 중 “DAUM 영화”를 선택한 이유는 각 영화에 대한 사용자 영화평점과 댓글의 수를 영화 개봉 전과 후로 제공해 시점별 자료수집이 가능하고 텍스트 마이닝을 위한 댓글의 내용을 읽는데 있어 “네이버 영화”와 같은 다른 사이트보다 용이하기 때문이다. 정보원천 변수에 대해 조금 더 자세하게 설명하면 우선 온라인 구전의 규모가 영화흥행에 미치는 영향을 파악하기 위해 각 영화에 대한 댓글의 수를 합산하여 변수로 사용했다. 또한 온라인 구전의 방향성이 영화흥행에 미치는 영향을 알아보기 위해 기존 연구에서 사용되었던 소비자 평균 평점을 변수로 설정했다. 이에 추가로 10점 척도로 측정되는 평점이 구전의 방향성을 측정하는데 있어 정확도가 떨어질 수 있다는 점을 보완하기 위해 오피니언 마이닝을 통해 분석한 각 댓글의 내용을 토대로 구전의 방향성을 파악해 변수로 사용했다. 마지막으로 영화 개봉 전과 후 정보원천 변수가 최종 영화흥행에 미치는 영향이 변화하는지에 대해 알아보기 위해 모든 정보원천 변수는 영화 개봉 전과 후에 작성된 온라인 구전 평점과 리뷰 자료를 각각 별도로 처리해 분석에 사용했다.

## 2. 분석 모형

### 2.1 구전 방향성 측정을 위한 오피니언 마이닝

본 연구는 영화흥행에 영향을 미치는 정보원천요소 중 하나인 소비자 오피니언을 더욱 상세히 측정하기 위해 소비자들의 기대감 혹은 경험이 서술된 댓글을 분석해 이와 관련된 변수를 정의한다. 이와 같은 오피니언 마이닝을 위한 텍스트 분류에는 기계학습방법과 같이 미리 알려진 분류결과를 바탕으로 학습을 실행한 후 새로운 자료의 분류결과를 알아보는 방식과 감성에 알맞게 구축된 사전에 문서의 단어를 매칭 시켜 분류하는 사전기반 분류가 있다. 본 연구에서는 영화와 관련된 댓글을 구성하는 단어를 기반으로 감성사전을 구축한 후 댓글의 내용을 극성분류하는 사전기반 분류를 사용했다.

우선 감성 단어사전 구축을 위해 “DAUM 영화”에서 2010년 11월부터 2015년 5월까지 국내에서 개봉한 모든 영화의 89,491개의 댓글을 수집하였다. 감성 단어사전으로 분류하기 이전 사전처리로 댓글의 기호, 조사, 숫자와 같이 분류에 필요없는 단어를 제거했다. 보통 극성에 대한 원인은 동사, 형용사, 일반명사에서 나타나는 바, ‘꼬꼬마 형태소 분석기(<http://kkma.snu.ac.kr/documents/index.jsp>)’를 사용하여 품사를 나눈 뒤 동사, 형용사, 일반명사를 수집해 1음절(Unigram) 단위의 단어를 긍정·부정·중립으로 구분하였다. 분석에 사용된 감성 단어사전은 각각 1360개, 2236개, 19480개의 긍정·부정·중립단어로 구성되었다.

최종적으로 본 연구의 분석대상인 537개의 영화와 관련된 59044개의 댓글에서 사용된 단어와 위에 구축한 감성 단어사전의 단어를 비교하여 각 댓글의 극성지수를 계산한 다음 긍정·부정·중립 댓글로 분류했다. N개의 분류가능한 단어로 구성된 댓글 j의 극성지수는 식(1)로 계산할 수 있다.

$$\text{극성지수}_j = \frac{\sum_{i=1}^N \text{긍정매칭}_i - \sum_{i=1}^N \text{부정매칭}_i}{\sum_{i=1}^N \text{긍정매칭}_i + \sum_{i=1}^N \text{부정매칭}_i} \quad (1)$$

식(1)에서 볼 수 있듯이 각 댓글의 극성지수는 1과 -1사이의 수로 표현이 되며 양수일 경우 긍정적 댓글, 음수일 경우 부정적 댓글 그리고 0일 경우 중립적 댓글로 분류된다. 구체적인 예를보면, 지난 2014년 개봉한 영화 ‘명량’의 게시판에 작성된 댓글인 “아들이 꼭 보고싶어 하는 영화여서 같이 봤는데 저는 그럭저럭. 아들은 재미있다고 하네요”를 살펴보면, 긍정을 표현하는 ‘보고싶어’, ‘재미있다고’ 단어와 부정을 표현하는 ‘그럭저럭’과 중립 혹은 특정 감성을 표현하지 않는 단어들로 구성되어 있다. 이 댓글의 극성지수를 식(1)에 따라 계산하면 총 3개의 유의미한 단어 가운데 2개의 긍정단어와 1개의 부정단어가 존재하므로 극성지수는 0.33이 되며 이 댓글은 긍정적 댓글로 분류가 된다. 이와 반대로 동 영화와 관련된 “좋은 재료를 가지고 맛있는 음식을 만들었네요. 도저히 이해불가”와 같은 경우, ‘좋은’이라는 1개의 긍정단어와 ‘맛없는’, ‘도저히’, ‘이해불가’라는 3개의 부정단어로 구성되어 있다. 따라서 이 댓글의 극성지수는 -0.5가 되며 부정적 댓글로 분류가 된다.

## 2.2 영화흥행요인 분석모형

본 연구에서는 영화흥행요인을 제작자 측면의 내재적 요소와 소비자 측면의 정보원천 요소를 분석하고, 뿐만 아니라 더욱더 구체적으로 정보원천 변수의 시점별 영향을 알아보기 위해 모형을 두 단계로 나누어 분석하였다. 우선 영화개봉 전 소비자들의 기대감에 따른 온라인 구전효

과가 영화 흥행에 미치는 효과를 파악하기 위해 다중회귀분석 모형을 구성했다. 종속 변수로는 앞서 설명했듯이 개별 영화의 최종 매출액을 로그 변환한 이후 사용했으며, 영화 흥행의 내재적 요인과 관련된 통제변수로는 개봉 스크린 수, 감독과 주연배우 2명의 영향력, 배급사 등급, 개봉 시즌, 장르, 관람등급 그리고 국내 제작 여부를 사용했다. 정보원천 변수로는 기존 연구에서 흔히 사용된 소비자들의 평균 평점으로 온라인 구전의 방향성을, 리뷰의 수를 온라인 구전의 규모 변수로 포함시켰는데 각 영화의 개봉 시점 이전까지 “DAUM 영화”에 소비자들이 작성한 3465개의 댓글과 평점을 바탕으로 변수를 구했다. 이 분석모형을 회귀식으로 표현하면 식(2)와 같다.

이후 추가적으로 정보원천 요인을 더욱 자세히 파악하기 위해 각 영화의 댓글을 오피니언 마이닝을 통하여 긍정 댓글과 부정 댓글로 구분한 다음 분석을 진행했다. 이것은 기존의 연구들에서 진행된 온라인 구전에 대한 측정을 온라인 평점과 단순 리뷰의 규모로서 효과를 평가하여 리뷰의 내용적인 부분에 대한 효과적인 측정이 제기되는 부분을 보완한 것이다. 이는 단순 댓글의 수로 온라인 구전 규모를 정의한 기존 연구들과는 달리 긍정적 온라인 구전과 부정적 온라인 구전의 규모를 파악함으로써 규모와 방향성의 복합적인 영향력을 알 수 있고, 주로 10점 척도로 평가되는 온라인 평점과 리뷰의 규모로서 측정하는 온라인 구전 효과에 단순함을 보완할 수 있는 장점이 있다. 단 원천적으로 긍정 댓글과 부정 댓글의 수가 총 댓글의 규모에서 나오는 파생변수라는 점을 고려해 총 댓글 중 긍정댓글의 비율을 계산한 후 이를 식(3)에 추가했다. 조금 더 구체적으로 식(2)와 식(3)의 차이는 소비자 오피니언을 측정하는 긍정댓글 변수의 유무로 설명할 수 있다. 이러한 전처리 과정을 거친 후 모든 주요 변수의 VIF

값은 10 이하로, 다중공선성 문제는 존재하지 않아 분석에 적합했다.

$$\begin{aligned} \log(sales_i) = & \beta_0 + \beta_1 Scr_i + \beta_2 dPow_i + \beta_3 act1Pow_i \quad (2) \\ & + \beta_4 act2Pow_i + \beta_5 distA_i + \beta_6 Season_i + \beta_7 Action_i \\ & + \beta_8 Comedy_i + \beta_9 DP_i + \beta_{10} DP18_i + \beta_{11} national_i \\ & + \beta_{12} AvgbefReview_i + \beta_{13} befReviewVol_i + \epsilon_i \end{aligned}$$

$$\begin{aligned} \log(sales_i) = & \beta_0 + \beta_1 Scr_i + \beta_2 dPow_i + \beta_3 act1Pow_i \quad (3) \\ & + \beta_4 act2Pow_i + \beta_5 distA_i + \beta_6 Season_i + \beta_7 Action_i \\ & + \beta_8 Comedy_i + \beta_9 DP_i + \beta_{10} DP18_i + \beta_{11} national_i \\ & + \beta_{12} AvgbefReview_i + \beta_{13} befReviewVol_i \\ & + \beta_{14} befReviewPosRatio_i + \epsilon_i \end{aligned}$$

다음 단계로 개봉 후 온라인 구전과 영화 흥행의 관계를 살펴보기 위해 앞선 모형 (2)와 (3)에 사용되었던 온라인 구전 변수를 개봉 후 데이터를 통해 얻은 변수로 교체했다. 개봉 전 온라인 구전은 주로 영화에 대한 소비자들의 기대감이 주로 나타남에 반해 개봉 후 온라인 구전은 아직 영화를 관람하지 않은 소비자들의 기대감도 있지만, 상당히 많은 관람객들의 경험과 개별적 평가가 동시에 반영되었다고 볼 수 있다. 영화 흥행과 관련해 통제변수로 사용된 모든 내재적 변수는 영상 상영기간동안 급변하지 않는다는 가정 아래 기존 모형과 동일하게 사용했다. 이에 따라 개봉 후 온라인 구전효과를 분석한 식은 식(4)~(5)와 같다.

$$\begin{aligned} \log(sales_i) = & \beta_0 + \beta_1 Scr_i + \beta_2 dPow_i + \beta_3 act1Pow_i \quad (4) \\ & + \beta_4 act2Pow_i + \beta_5 distA_i + \beta_6 Season_i + \beta_7 Action_i \\ & + \beta_8 Comedy_i + \beta_9 DP_i + \beta_{10} DP18_i + \beta_{11} national_i \\ & + \beta_{12} AvgAftReview_i + \beta_{13} AftReviewVol_i + \epsilon_i \end{aligned}$$

&lt;표 4&gt; 분석 결과(종속 변수 : 로그매출액)

	모형(1)		모형(2)		모형(3)		모형(4)	
	계수	VIF	계수	VIF	계수	VIF	계수	VIF
절편	19.758** (0.492)		21.190** (0.669)		18.596** (0.477)		20.416** (0.673)	
<i>Scr</i>	0.0034** (0.0006)	1.112	0.0033** (0.0006)	1.112	0.0033** (0.0006)	1.114	0.0032** (0.0006)	1.114
<i>dPow</i>	0.0026 (0.011)	1.156	0.0021 (0.011)	1.157	0.0062 (0.0109)	1.148	0.0046 (0.0108)	1.150
<i>act1Pow</i>	0.0190* (0.0095)	1.184	0.0179* (0.0089)	1.186	0.0198* (0.0093)	1.182	0.0194* (0.0092)	1.182
<i>act2Pow</i>	0.0174* (0.0084)	1.219	0.0197* (0.0083)	1.229	0.0148* (0.007)	1.219	0.0174* (0.0082)	1.229
<i>distA</i>	1.062** (0.161)	1.086	1.091** (0.159)	1.090	0.994** (0.157)	1.087	1.101** (0.1557)	1.087
<i>Season</i>	0.304* (0.138)	1.028	0.292* (0.137)	1.029	0.292* (0.135)	1.029	0.279* (0.133)	1.029
<i>Action</i>	0.459** (0.167)	1.329	0.423* (0.167)	1.336	0.521** (0.165)	1.329	0.486** (0.163)	1.334
<i>Comedy</i>	-0.346 (0.198)	1.171	-0.288 (0.197)	1.182	-0.307 (0.195)	1.173	-0.224 (0.193)	1.188
<i>DG</i>	-0.606** (0.188)	1.414	-0.566** (0.187)	1.420	-0.753** (0.186)	1.433	-0.724** (0.184)	1.435
<i>DG18</i>	-0.910** (0.177)	1.131	-0.916** (0.176)	1.131	-0.823** (0.175)	1.147	-0.822** (0.174)	1.147
<i>national</i>	-0.971** (0.169)	1.352	-0.998** (0.167)	1.356	-1.013** (0.165)	1.352	-1.035** (0.164)	1.353
<i>AvgbefReview</i>	0.079 (0.087)	1.047	0.1021 (0.089)	1.204				
<i>AvgAftReview</i>					0.2651** (0.054)	1.127	0.3115** (0.0549)	1.186
<i>befReview Vol</i>	0.0098** (0.0027)	1.075	0.0093** (0.0027)	1.078				
<i>AftReview Vol</i>					0.0004** (0.0001)	1.089	0.0004** (0.0001)	1.093
<i>befReviewPosRatio</i>			2.187** (0.699)	1.109				
<i>AftReviewPosRatio</i>							2.471** (0.653)	1.111
모형 F 검정값	23.52**		22.91**		25.6**		25.4**	
<i>AdjR</i> <sup>2</sup>	0.343		0.365		0.373		0.389	

\* :  $p < 0.05$ , \*\* :  $p < 0.01$  ;

$$\begin{aligned} \log(\text{sales}_i) = & \beta_0 + \beta_1 \text{Scr}_i + \beta_2 \text{dPow}_i + \beta_3 \text{act1Pow}_i \quad (5) \\ & + \beta_4 \text{act2Pow}_i + \beta_5 \text{distA}_i + \beta_6 \text{Season}_i + \beta_7 \text{Action}_i \\ & + \beta_8 \text{Comedy}_i + \beta_9 \text{DP}_i + \beta_{10} \text{DP18}_i + \beta_{11} \text{national}_i \\ & + \beta_{12} \text{AvgbefReview}_i + \beta_{13} \text{befReview Vol}_i \\ & + \beta_{14} \text{AftReviewPosRatio}_i + \epsilon_i \end{aligned}$$

최종적으로 분석에 사용된 모든 변수의 이름과 그 의미는 <표 2>와 같다.

## IV. 실증분석

### 1. 기술통계 분석

수집한 자료의 기초통계분석을 실시한 결과 <표 3> 몇몇 흥미로운 점을 발견했다. 우선 2012년 1월부터 2014년 12월 사이 개봉한 537개의 표본대상 중 32.5%인 175개 영화가 국내제작영화인 것으로 밝혀졌다. 개봉한 영화의 절대적인 수치는 해외제작영화이며, 이중 90% 이상이 미국에서 제작된 영화이다. 또한 해외 제작영화는 국내제작영화의 2배 정도임에도 불구하고 영화의 평균 매출액은 국내제작영화가 146억 원으로 해외제작영화의 평균 매출액을(54억 원) 크게 상회하는 것으로 나타났다. 이는 지난 수년간 국내영화의 질적 향상을 간접적으로 보여줄 뿐만 아니라 한국 소비자들의 취향을 잘 반영하였다고 추정된다. 이를 통해 한국영화시장 내에서 국내제작 영화의 경쟁력이 우수하다는 점을 시사한다. 한편 국내영화시장에서 배급사가 차지하는 영향력을 살펴보면 상위 20%에 속하는 A급 배급사가 수집 기간 중 개봉한 537개 영화 가운데 74%에 달하는 395개 영화를 배급하여 심각하게 편중되어 있다. 또한, 이러한 A급 배급사들이 배급한 영화의 평

균 매출액은 102억 원으로 그 외 배급사들이 배급한 영화의 3~4배에 이르는 평균 매출액을 거두었다. 이는 거대 배급사와 상영관 사이에 긴밀한 협력구조를 가지고 있는 한국 영화시장의 특성을 나타내며, 거대 배급사가 배급하는 일부 영화들의 과도한 스크린 독점이 발생한다고 볼 수 있다. 마지막으로 온라인 구전효과의 방향성을 보여주는 평점을 살펴보면, 개봉 전 평점의 평균은 8.42인 것에 비해 개봉 후 평점의 평균은 7.30으로 1점 이상의 유의미한( $t$ -검정,  $\alpha = 0.01$  유의수준) 차이를 보인다. 특히 개봉 전 평점과 같은 경우 표준편차가 불과 1.46으로 변동성 또한 개봉 후 평점보다(표준편차 3.24) 현저히 낮은 수준임을 알 수 있다. 흥미롭게도 이러한 현상은 개봉 전 평점이 일반적으로 높은 점수대에 형성되며 개별 영화간 평점의 차이도 그다지 크지 않다는 점을 시사한다. 이와 같이 평점은 작성 시점별로 다른 특성을 보이며, 각 시점에 따른 온라인 구전효과를 별개로 분석해야 할 필요성을 보인다. 이에 반해 개봉 전/후의 긍정댓글 비율은 각각 0.77과 0.80으로 비슷한 수준이며 표준편차 또한 크게 차이를 보이지 않는다. 이와 같은 변수 속성의 일관성은 영화의 평점보다 긍정댓글의 비율이 온라인 구전효과의 방향성을 측정하는데 있어 더욱 효율적일 가능성을 보여준다.

### 2. 분석결과

연구모형을 실증분석한 결과는 <표 4>에서 확인할 수 있다. 회귀식 (2)~(3)을 적용한 모형 (1)~(2)는 영화가 개봉하기 이전 온라인구전의 효과를 알아본 모형이며 회귀식 (4)~(5)를 적용한 모형 (3)~(4)는 영화 개봉 이후 온라인구전 효과를 분석한 모형이다.

&lt;표 2&gt; 온라인구전과 영화흥행의 관계분석에 사용된 변수 항목

변수명	의미
$\log(sales_i)$	영화 $i$ 총 매출의 로그변환값
$Scr_i$	영화 $i$ 의 개봉 스크린 수
$dPow_i$	영화 $i$ 의 감독이 개봉 시점 이전 5년간 감독한 영화 수
$act1Pow_i$	영화 $i$ 의 첫번째 주연배우가 이전 5년간 출연한 영화 수
$act2Pow_i$	영화 $i$ 의 두번째 주연배우가 이전 5년간 출연한 영화 수
$distA_i$	영화 $i$ 배급사의 등급이 상위 20%인 경우 1
$Season_i$	영화 $i$ 개봉시기가 성수기일 경우 1
$Action_i$	영화 $i$ 의 장르가 액션일 경우 1
$Comedy_i$	영화 $i$ 의 장르가 코미디일 경우 1
$DG_i$	영화 $i$ 관람등급이 전체관람가일 경우 1
$DG18_i$	영화 $i$ 관람등급이 청소년관람불가일 경우 1
$national_i$	영화 $i$ 의 제작국가가 외국인 경우 1
$AvgbefReview_i$	영화 $i$ 의 개봉 이전 평균 평점
$AvgAftReview_i$	영화 $i$ 의 개봉 이후 평균 평점
$befReviewVol_i$	영화 $i$ 의 개봉 이전 작성된 댓글의 갯수
$AftReviewVol_i$	영화 $i$ 의 개봉 이후 작성된 댓글의 갯수
$befReviewPosRatio_i$	영화 $i$ 의 개봉 이전 작성된 댓글 중 긍정댓글 비율
$AftReviewPosRatio_i$	영화 $i$ 의 개봉 이후 작성된 댓글 중 긍정댓글 비율

&lt;표 3&gt; 수집 표본에 대한 기초통계량 분석

	평균	중앙값	표준편차	관측 수
매출액(백만원)	8377	2122	1581	537
로그매출액	21.37	21.48	1.95	537
개봉관 수	192.8	186	111.3	537
감독 영향력	7.89	6	6.53	537
주연배우1 영향력	13.33	13	7.21	537
주연배우2 영향력	12.91	12	8.92	537
A급 배급사 배급 영화				395
성수기 개봉 영화				275
액션 장르 영화				173
코미디 장르 영화				88
전체관람가 영화				131
청소년관람불가 영화				112
국내제작 영화				175
외국제작 영화				362
개봉 전 평점	8.42	8.70	1.46	537
개봉 후 평점	7.30	7.50	3.24	537
개봉 전 댓글 수	6.45	4	22.82	3465
개봉 후 댓글 수	103	35	416.9	55579
개봉 전 긍정댓글 비율	0.77	0.71	0.112	537
개봉 후 긍정댓글 비율	0.80	0.75	0.106	537

## 2.1 통제변수 결과 분석

본 연구에서 분석한 다중회귀모형 (1)~(4)의 전반적 모형적합성 여부를 F-검정을 통해 살펴본 결과 모든 모형이 적합한 모형인( $p < 0.01$ ) 것으로 확인되었다. 우선 통제변수로 사용한 제작자 측면의 내재적변수들과 관련된 다중회귀분석모형 (1)~(4)의 결과를 살펴보면 대체로 기존 연구와 동일한 결과를 보인다. 우선 스크린과 A급 배급사 변수의 계수를 살펴보면 둘 다 유의한 양(+)의 값을 보여주고 있는데 이는 한국영화 흥행 결정요인에 대한 김은미(2003)의 연구의 결과와 동일하다. 위에서 언급한 바와 같이 국내 영화산업의 특성상 일부 대규모 배급사들이 영화의 투자도 동시에 진행해 다수 영화의 배급권리를 가져오며, 또한 상영관과 긴밀한 관계를 통해 스크린을 우선 확보 할 수 있다. 물리적으로 한정되어 있는 상영관을 특정 배급사들의 영화로 독점 할 시 관람객들의 선택은 제한되고, 이에 따라 이러한 영화들의 흥행 가능성은 높아질 수밖에 없다는 결론을 본 연구에서도 확인할 수 있다. 둘째로, 영화의 '질'과 연관되어 있는 인적요소 변수 중 배우 영향력 변수는 기존의 연구들(유현석 2002; 김은미 2003)에서 밝혀진 바와 같이 영화 흥행과 유의한 정의 관계를 모든 다중회귀분석 모형에서 보였다. 이를 통해서 배우들이 실질적으로 영화 매출에 매우 중요한 요인으로 보여지며, 아울러 소비자들이 영화를 선택할 때 배우가 중요한 기준이 되는 것으로 추정된다. 한편 감독 영향력 변수는 모든 회귀식에서 별다른 영향을 미치지 못하는 것으로 나타났다. 배우들은 소비자와 스크린에서 직접적으로 만나거나 또는 다른 여러 매체를 통해서 인지를 알리는 경우가 많지만, 감독은 실질적으로 스크린에서 노출이 되는 경우가 거의 없고, 하나의 작품에 집중하는 성향을 보여 작품의 편수도 적으며 작품과 작품사이의 휴식기도 배우에 비해 긴 것이

감독 영향력과 관련된 결과의 원인으로 추정된다. 또한 최근 들어 여러 신인 감독들이 선전하는 현상이 자주 발생하는 것을 보아, 소비자들의 입장에선 해당 작품의 감독보다는 더 잘 알려진 배우들의 출연이 관람여부를 결정하는데 크게 작용할 수 있다. 개봉시즌 또한 영화 흥행에 영향을 미치는 것으로 확인되었다. 모든 모형에서 성수기 시즌에 개봉한 영화가 비수기 시즌에 개봉한 영화에 비해 평균 매출이 유의하게 큰 것으로 나타났으며 이는 김휴종(1998)의 연구결과와 동일하다. 장르 변수와 같은 경우 코미디 장르는 영화 흥행과 유의한 관계가 없는 것으로 나타난 반면 액션 장르는 영화 흥행에 유의한 정의 영향을 미치는 것으로 나타났다. 그러나 장르 변수와 같은 경우 기존 연구들마다 영화 흥행에 유의한 영향을 미치는 장르가 다르게 나타나는 것으로 미루어 연구시점에 수집된 흥행영화들 중 해당 장르의 영화가 흥행을 한 것이지 일반적으로 일정 장르의 영화가 항상 흥행에 유의한 영향을 미친다고 결론을 내리기는 어렵다. 관람등급과 관련하여 전체 관람가 영화와 청소년관람불가 등급의 영화가 그 중간 등급들의 영화와 비교해 유의한 부의 관계를 가지는 것으로 나타났다. 이는 관람등급과 영화 흥행이 역 "U"자 관계를 지닌다는 기존 연구(배정호 심범준 김병도 2010)의 결과와 유사하다. 다시 말해, 영화의 내용적인 측면을 간과한 하고, 오로지 관람등급 요소인 연령이라는 기준을 통해서 영화입장을 제한함으로써 영화 흥행에 부정적으로 영향을 미치는 부분이 있는 것으로 보인다. 마지막으로 해당 기간 동안 수집된 국내 개봉 영화 가운데 국내 제작 영화가 해외 제작 영화보다 더 큰 흥행성적을 거둔다는 점을 볼 수 있다. 과거 연구(Waterman and Jayakar, 2000)는 제작비 규모에서 비교하기 어려울 만큼 큰 차이를 보이는 미국 제작 영화의 점유율이 전세계적으로 증가 추세에 있고 미국 영화시장의 규모에서 발생하는 경쟁적인 이점을 통해 유



통채널을 독점한다고 주장했다. 그러나 본 연구 결과에 따르면 영화매출이라는 측면에서 국내제작영화는 해외제작영화와 비교해 경쟁력이 있다. 이러한 결과는 단순히 제작비 규모가 크다고 해서 영화 흥행에 영향을 미치는 것은 아니라는 점을 시사한다. 또한 국내제작 영화가 많은 부분 국내 소비자들의 니즈(need)와 원츠(want)를 잘 반영한 것으로 보인다. 따라서 선행연구에서 연구된 바 있는 제작자 측면의 내재적 변수는 본 연구에서도 감독 변수만을 제외하고는 모두 영화매출에 영향을 준다는 것을 확인할 수 있다.

## 2.2 온라인 구전 방향성과 규모의 영향 분석

우선 온라인 구전효과와 관련된 변수들 중 모형 1의 개봉 전 온라인 평균 평점을 제외한 모든 변수가 영화흥행에 유의한 양의 영향을 미치는 것으로 나타났다. 온라인 구전효과의 포괄적인 영향이 존재하는지 여부를 통계적으로 검정하기 위해 온라인 구전효과 변수가 포함된 모형과 온라인 구전효과 변수를 제외하고 내재적 변수만 사용한 모형을 부분검정한 결과, 온라인 구전효과와 관련된 변수들이 영화매출에 대해 추가적으로 설명하는 부분은 유의미한 것으로 밝혀졌다(부분  $F$ -검정통계량: 16.878,  $\alpha = 0.01$  유의수준). 따라서 온라인 구전효과의 포괄적 영향을 살펴본 가설 1은 지지된다. 온라인 구전효과를 세부적으로 살펴보면, 첫째로 온라인 댓글의 규모는 분석에 사용된 모든 모형에서, 개봉 전/후 시점에 상관없이 영화매출과 유의한( $p < 0.01$ ) 정의 관계를 갖는 것으로 나타났다. 이는 영화산업에서 온라인 구전효과를 분석한 기존의 연구들(Chintagunta et al. 2006; Liu 2006; 배정호 심범준 김병도 2010)과 동일한 결과로 온라인에서 발생하는 구전의 규모가 클수록 높은 영화매출과 상관관계를 보인다는 점을 의미한다. 이러한 결과를 통해 본 연구는 온라인 댓글의 수가

시점에 상관없이 구전효과의 근본적인 강도를 나타내는 좋은 지표로 활용될 수 있음을 재차 확인했다. 결과적으로 온라인 구전효과의 규모가 커질수록 해당 영화가 대중들에게 노출되는 횟수가 많아진다. 또한 더 많은 관심을 끄는 영화일수록 더 많은 댓글로 그 관심도가 나타나며 그 시점이 영화의 개봉 전 이면 후건 영화매출에 영향을 주는 다른 잠정적 소비자들의 관람결정에 영향을 준다고 판단할 수 있다. 이러한 결과를 토대로 가설 2b는 지지된다.

구전의 방향성을 측정한 평균평점의 경우 개봉 전 시점의 평점을 살펴본 모형 다중회귀모형(1)~(2)에선 영화매출에 영향을 주지 못하는 반면, 모형(3)~(4)의 분석결과에서 볼 수 있듯이 개봉 후 시점의 평점은 영화매출과 유의한( $p < 0.01$ ) 정의 관계를 가진다. 따라서 가설 2a는 지지되지 못한다. 앞선 선행연구를 잠시 살펴보면, Liu(2006)와 Duan et al.(2008)과 같은 경우 구전의 방향성은 영화매출과 관계가 없다고 주장한 바에 비해 배정호, 심범준 김병도(2010)의 연구에서는 둘 사이의 관계가 존재한다고 주장했다. 특히 배정호, 심범준 김병도(2010)의 경우 이러한 연구결과의 차이가 분석 자료의 출처국가가 다르기 때문이라고 주장하며 좀 더 근본적으로는 북미 문화와 아시아 문화에 내포된 자아관점으로 나뉜다고 설명했다. 즉 독립적 자아관점을 가진 서양인들의 경우 타인의 평가에 대한 의존하는 정도가 상호의존적 자아관점을 지닌 동양인에 비해 작다는 것이다. 그러나 본 연구의 결과를 살펴보면 동일한 동양인들을 대상으로 표본을 추출했음에도 불구하고 평점을 작성한 시점이 개봉 전이냐 후이냐에 따라 온라인 구전 방향성의 효과가 달라진다는 것을 알 수 있다. 따라서 구전 방향성을 측정한 지표의 신뢰성에 근거하여 본 연구결과와 기존 연구결과의 차이점을 설명할 수 있다. 앞서 기술통계분석을 통해 알아본 바와 같이 개봉 전 평점은 영화에 상관없이 대체적으로 높은 점수에 집

중되어 있고 변동성 또한 부족해 변별력이 없는 지표라 평가할 수 있다. 이는 평가의 시점 상 해당 영화를 관람하지 않은 채 점수를 부여하기 때문에 나타나는 현상이라고 할 수 있다. 개봉 전 평점과 같은 경우 소비자가 이미 이러한 단점을 파악하고 있어 지표를 신뢰하지 않으며 이에 따라 영화매출에 대한 영향력이 없다고 설명할 수 있다. 반면, 개봉 후 평점은 잠재소비자들에게 실제 경험을 바탕으로 작성된 비교적 객관적 평가라는 믿음을 주는 신뢰도가 높은 지표이며, 따라서 영화흥행에도 긍정적인 영향을 준다. 이에 추가적으로 표본의 규모에 의한 평점의 신뢰도 문제를 생각해 볼 수 있다. 위에서 언급하였듯이 평균 평점이라는 지표는 참여 평가자의 수가 적을수록 극한값에 더 큰 영향을 받는다. 기술 통계를 살펴보면 평균적으로 개봉 전 영화에 대한 리뷰나 평점은 영화당 6.45개인 것에 비해 개봉 후에는 평균 103개인 것을 볼 수 있다. 이를 바탕으로 리뷰가 많은 영화의 경우 기존 연구(배정호, 심범준, 김병도 2010)와 같이 평균 평점은 구전효과의 방향성을 측정하는데 적절하다고 볼 수 있으나 리뷰가 많지 않은 영화의 구전 방향성은 평균 평점으로 파악하기가 어렵다는 점을 시사한다.

본 연구에서는 기존 연구에서 제시한 평균 평점과는 달리 온라인 구전 방향성의 새로운 지표로 긍정댓글비율을 제시했다. 긍정댓글비율 변수를 사용한 모형(2)와 모형(4)의 결과를 살펴보면 개봉 전/후 시점에 상관없이 긍정댓글비율과 영화매출은 항상 유의한( $p < 0.01$ ) 정의 관계를 보인다. 그러므로 가설 2c는 지지됨을 확인할 수 있다. 이는 해당 영화와 관련된 댓글 중 긍정적인 댓글이 많을수록 높은 영화매출과 관계있다는 점을 시사하고 결국 소비자 의사결정 프로세스의 한 축으로 다른 이의 평가 혹은 선호여부가 존재함을 알 수 있다. 또한 이 결과는 평균 평점으로 측정한 구전 방향성 결과와는 다른 특성을 보인다. 제한된 10점 척도 안에

서 단순히 부여된 숫자의 취합인 평균 평점을 통해 구전 방향성을 측정할 시 소비자들이 그 평점을 신뢰하느냐 혹은 충분한 표본을 바탕으로 평균 평점을 구했는지 여부에 따라 효과 또한 달랐다. 그러나 댓글의 내용은 소비자들이 직접 읽고 정성적으로 판단하기 때문에 댓글의 내용을 기반으로 작성된 긍정댓글비율은 편향성 문제에서 조금 더 자유로우며 표본의 크기가 작더라도 리뷰를 읽고 판단함으로써 다른 소비자들의 오피니언을 좀 더 정확히 알아볼 수 있다. 구전 방향성의 지표로 소비자오피니언 변수를 추가해 분석한 모형 (2), (4)와 평균평점만을 사용한 모형 (1), (3)을 비교하여 부분검정을 실시한 결과, 개봉 전 온라인 구전효과를 분석한 모형 (1)~(2)에서 소비자오피니언 변수는 영화매출에 대해 유의미한 추가 설명력을 지닌 것으로 나타났으며 (부분  $F$ -검정통계량: 9.78,  $\alpha = 0.01$  유의수준) 개봉 후 모형에서도 소비자오피니언 변수는 유의한 설명력을 지닌 변수로 확인되었다(부분  $F$ -검정통계량: 14.332,  $\alpha = 0.01$  유의수준).

## V. 결론

### 1. 연구의 요약

본 연구에서는 소비자 오피니언이 영화 흥행에 영향력을 미치는 부분을 오피니언 마이닝 기법을 통해 분석하였다. 특히 선행연구에서 활발히 연구된 바 있는 영화 제작자 측면의 내재적 변수와 본 연구에서 알아보고자 했던 소비자오피니언을 반영한 정보원천 변수로 나누어서 복합적인 분석을 진행하였다. 우선 선행연구에서 연구된 스크린 수, 배우, 감독, 배급사, 개봉시즌, 장르, 관람등급 그리고 제작국가와 같은 영화흥행의 내재적 변수와 영화매출과의 관계는 기존의 선행연구와

비슷한 결과가 나타났다. 내재적 변수 중 감독 영향력 변수만 영화매출에 별다른 영향력이 없는 것으로 나왔고, 스크린 수, 배우, 배급사, 개봉시즌, 장르, 관람등급 그리고 제작 국가는 모든 다중회귀모형식에서 영화매출과 유의미한 양의(+) 관계가 있는 것으로 분석되었다.

소비자오피니언을 반영한 정보원천 변수인 온라인 구전효과와 영화매출 사이의 관계를 분석한 결과에서는 몇 가지 흥미로운 점을 발견할 수 있었다. 우선 본 연구에서는 선행연구와 같이 온라인 구전효과를 크게 온라인 구전의 방향성과 규모로 나누어 각각 영화홍행에 어떠한 영향을 미치는지 알아보았다. 여기에 기존 선행연구에서는 사용되지 않았던 오피니언 마이닝 기법을 활용함으로써 온라인 구전 방향성의 효과를 더욱 자세하게 파악하는 한편, 구전효과를 영화개봉전과 개봉이후로 나누어서 시점에 따른 구전효과의 변화를 분석하였다. 온라인 구전효과와 관련해 첫째, 온라인 구전의 규모는 영화개봉 전이나 개봉 이후 모두 영화홍행에 유의한 정의 영향을 미치는 것으로 나타났다. 둘째, 평균 평점으로 측정된 구전의 방향성의 경우 영화개봉 이전의 구전 방향성은 영화홍행에 별다른 영향력이 없으나, 영화개봉 이후에는 유의미한 관계가 있는 것으로 나왔다. 마지막으로 긍정적인 리뷰의 비율로 측정한 소비자 오피니언은 개봉 전의 오피니언이나 개봉 후의 오피니언이나 모두 영화 매출에 유의미하게 영향을 주는 것으로 분석된다. 전반적으로 온라인 구전의 효과는 방향성(Valence)과 규모(Volume) 모두 영향을 주는 것으로 보여지며, 영화 개봉이후에는 특히 그 효과가 아주 큰 것으로 분석된다.

## 2. 연구의 시사점

이러한 결과는 영화산업 관계자들이 영화홍행을 위한 전략을 수립하는데 있어 포괄적인 방법을 제시한다. 우선적으로 사전 기획단계에서는 기존 선행연구에서도 밝혀진 바 있는 내재적변수의 영향력에 집중하여 영화제작과 관련된 의사결정을 내릴 필요가 있다. 영화의 완성이 다가오는 시점에서는 본 연구에서 살펴본 온라인 구전효과를 바탕으로 전반적인 영화 홍보 전략을 세우는 한편 소비자들의 반응을 지속적으로 모니터링해 부족한 부분을 집중적으로 공략하는 등 효과적인 마케팅 전략을 수행할 수 있다. 예를 들어 개봉 전 리뷰의 규모를 파악하고 리뷰의 내용을 분석함으로써 구전의 방향성을 측정하여 영화홍행을 예측해 볼 수 있다. 또한 오피니언 마이닝 기법은 단어 사전의 정의에 공을 들여 데이터베이스를 신중히 구축할 경우 많은 리뷰 내용의 분석을 단시간에 마칠 수 있으며 사람이 직접 리뷰를 읽고 평가하는 방식보다 시간과 비용을 절감함으로써 제한된 마케팅 비용을 효율적으로 사용할 수 있다는 장점이 있다. 덧붙여 이러한 분석 프로세스를 정착시킨다면 정기적으로 새롭게 유입되는 리뷰를 분석함으로써 구전 방향성의 추이를 살펴보고 상황에 맞게 유연하게 대처할 수 있으며 간단한 단어 사전의 업데이트를 통해 적은 비용으로 시스템의 변화를 이룰 수 있다. 마지막으로 본 연구의 결과, 영화와 관련한 정보의 비대칭적 상태에서 일정 정도의 정보 제공은 긍정적으로 영화매출에 영향을 줄 수 있을 것으로 추정된다. 따라서 영화산업 관계자들은 여러 포털 사이트와의 협력 혹은 독자적인 플랫폼 개발을 통해 소비자들이 영화에 대해 토론하고 평가할 수 있는 장을 제공하고 고려하는 현실적 방안을 고민함으로써 온라인 구전

효과의 효율성을 극대화 시킬 수 있는 방법도 고려해야 할 필요가 있다고 판단된다.

### 3. 연구의 한계점 및 향후 연구

영화매출에 관계된 본 연구는 추후 몇 가지 측면을 더 고려한 후속 연구가 있어야 할 것으로 보인다. 우선 영화매출에 영향을 미치는 내재적 변수 중 주요 변수인 제작비를 고려할 필요가 있다. 특히 영화의 순수 제작비와 단계별 홍보비용 등 세부 예산의 지출과 관련된 자료가 추가된다면 여러 홍보 활동 중 효과적인 홍보활동은 무엇이며 예산 대비 온라인 구전효과의 실효성을 더욱 자세히 평가할 수 있을 것이다. 현재 영화의 제작비 같은 경우 대략적인 추정치 외 정확한 세부 예산은 공개하지 않고 있으며 대략적인 총제작비 또한 일부 영화만 공개된 상황이나 추후 세부 제작비와 관련된 자료를 구할 수 있다면 영화매출에 대한 연구에 많은 도움이 될 것으로 보인다.

또한 영화산업의 특성을 자세히 살펴보면 대개 주 단위로 어떠한 영화를 상영할지 몇 개의 스크린에서 상영할 것인지와 같은 다양한 의사결정을 지속적으로 내린다. 또한 이러한 의사결정은 해당 주의 영화흥행 여부에 따라 결정되는 등 본질적으로 영화 개봉 이후에는 영화흥행과 여러 흥행요인들 사이에 동적 관계가 있다고 볼 수 있다. 이러한 동적 관계는 스크린 수, 상영기간과 같은 내재적 변수 이외에 구전효과와 영화흥행의 관계에서도 존재할 것으로 추정된다. 따라서 향후 연구에서는 영화의 주별 관객 수, 주별 개봉관 수, 주별 정보원천 자료를 사용한 동적 분석을 통해 여러 흥행요인의 영향력을 시점별로 분석함으로써 그 효과를 더 상세히 파악할 필요가 있다.

추가적으로 인구통계학적인 자료를 포함한 분석이 가능하다면 좀 더 영화관계자들에게 실질적

인 도움이 될 것으로 판단된다. 이는 사람들의 나이, 성별, 직업, 결혼유무 그리고 소득 등이 어떤 영화선택과 관계가 있는지를 파악하고 영화 배급 단계에서부터 개별 소비자의 상황에 알맞은 영화 프로모션을 개인의 스마트기기를 통해 제공함으로써 효율적인 마케팅 전략을 실행할 수 있기 때문이다.

마지막으로 정보원천 변수들과 관련해 평점과 리뷰를 일반인리뷰와 전문가리뷰로 구분하여 각 주체별 구전의 효과가 동일하게 나타나는지를 비교해 볼 필요가 있다. 예를 들어, 개봉 전에는 전문가들의 구전효과가 초기 영화를 선택하는 소비자들에게 일반인들의 평가보다 더 큰 직접적인 영향을 미칠 수 있을 것으로 생각해 볼 수 있으며 이 부분에 대한 추가적인 연구가 필요할 것으로 판단된다.

논문접수일: 2017. 02. 20.

1차 수정본 접수일: 2017. 03. 21.

2차 수정본 접수일: 2017. 04. 14.

게재확정일: 2017. 04. 14.

### 참고문헌

- Ahn, Sungah and Kim, Taejoon (2003), "The Determinants of Opening Share and Decay Rate in Motion Pictures," *Korea Marketing Review*, 18(3), 1-17.
- Ainslie, Andrew, Drèze, Xavier and Fred Zufryden (2005), "Modeling Movie Life Cycles and Market Share," *Marketing Science*, 24(3), 508 - 517.
- Bae Jungho, Kim Byugdo and Shim Bumjun (2010), "Simultaneous Effect between eWOM and

- Revenues: Korea Movie Industry,” *ASIA Marketing Journal*, 12(2), 1-25.
- Chen, Hsinchun and David Zimbra (2010), “AI and Opinion Mining,” *Browse Journals and Magazines*, 25(3), 74-80.
- Chevalier, Judith A. and Dina Mayzlin (2006), “The effect of word of mouth on sales: Online book reviews,” *Journal of Marketing Research*, 43(3), 345-354.
- Chintagunta, Pradeep K., Gopinath, Shyam and Sriram Venkataraman (2010), “The Effects of Online User Reviews on Movie Box Office Performance: Accounting for Sequential Rollout and Aggregation Across Local Markets,” *Marketing Science*, 29(5), 944-957.
- Chintagunta, Pradeep K., Puneet Manchanda, Jean-Pierre Dubé and Khim Yong Goh (2006), “The Effect of Banner Advertising on Internet Purchasing,” *Journal of Marketing Research*, 43(1), 98-108.
- Dellarocas, Chrysanthos, Zhang, Xiaoquan (Michael) and Neveen F. Awad (2007) “Exploring the value of online product reviews in forecasting sales: The case of motion pictures.” *Journal of Interactive Marketing*, 21(4), 23 – 45.
- De Vany, A. and W. D. Walls, (1999), “Uncertainty in the Movie Industry: Does Star Power Reduce the Terror of the Box Office?,” *Journal of Cultural Economics*, 23(4), 285 – 318.
- Duan, Wenjing, Gu, Bin and Andrew B. Whinston (2008), “The dynamics of online word-of-mouth and product sales—An empirical investigation of the movie industry,” *Journal of Retailing*, 84(2), 233 – 242.
- Elberse, Anita, and Jehoshua Eliashberg (2003), “Demand and Supply Dynamics for Sequentially Released Products in International Markets: The Case of Motion Pictures,” *Marketing Science*, 22(3), 329 – 354.
- Fossen, Beth L. and David A. Schweidel (2017), “Television Advertising and Online Word-of-Mouth: An Empirical Investigation of Social TV Activity,” *Marketing Science*, 36(1), 105-123.
- Garrison, Lee Cedric, Jr (1971), “Decision Processes in Motion Picture Production: A Study of Uncertainty,” *Doctoral Dissertation*, Stanford University.
- Godes, David and Dina Mayzlin(2004), “Using Online Conversations to Study Word-of-Mouth Communication,” *Marketing Science*, 23(4), 545 - 560.
- Guan, Sunju (2014), “The Impact of Critics on Movie Market Performances: Art versus Commercial ,” *Cultural and Economic Research*, 17(3), 3-21.
- Jayakar, Krishna P. and David Waterman (2000), “The Economics of American Theatrical Movie Exports: An Empirical Analysis,” *Journal of Media Economics*, 13(3), 153-169.
- Khare, Adwait, Labrecque, Lauren I. and Anthony K. Asare (2011), “The Assimilative and Contrastive Effects of Word-of-Mouth Volume: An Experimental Examination of Online Consumer Ratings,” *Journal of Retailing*, 81(1), 111 – 126.
- Kim, Eunmee (2003), “The Determinants of Motion Picture Box Office Performance: Evidence from Movies Exhibited in Korea,” *Korean Journal of Journalism and Communication Studies*, 47(2), 190-220.
- Kim, Hyujong (1998), “A star movie star’s star power analysis,” *Cultural and Economic Research*, 1(1), 165-200.

- Kramer, Thomas, Spolter-Weisfeld, Suri and Maneesh Thakkar (2007), "The Effect of Cultural Orientation on Consumer Responses to Personalization," *Marketing Science*, 26(2), 246-258.
- Lee, Yanghwan, Chang, Byenghee and Park, Kyungwoo (2007), "An Exploratory Study for Comparing Factors Affecting Box Office Performances between Countries: Focusing on Performances of U.S. Movies in South Korea and U.S.," *Journal of Journalism and Communication Studies*, 7(1), 185-222.
- Litman, Barry R. (1983), "Predicting Success of Theatrical Movies: An Empirical Study," *Journal of Popular Culture*, 16(4), 159-175.
- Litman, Barry R. and Richard Ducey (1980), "Review: Broadcasting in the United States: Innovative Challenge and Organizational Control, Mosco Vincent," *Public Opinion Quarterly*, 44(4), 602-604.
- Litman, Barry R. and Linda S. Kohl (1989), "Predicting financial success of motion pictures: The '80s experience," *Journal of Media Economics*, 2(2), 35-50.
- Liu (2006), "Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue," *Journal of Marketing*, 70(3), 74-89.
- Liu, Yong, Anindita, Chakravarty, and Tridib Mazumdar (2010), "The Differential Effects of Online Word-of-Mouth and Critics' Reviews on Pre-release Movie Evaluation," *Journal of Interactive Marketing*, 24(3), 185-197.
- Neelamegham, Ramya and Pradeep Chintagunta (1999), "A Bayesian Model to Forecast New Product Performance in Domestic and International Markets," *Marketing Science*, 18(2), 115-136.
- Park, Hyunghyun and Park, Chansu (2001), "The Relationship between Critical Reviews and Performance of Movies : Does it Hold in the Internet Era?," *Korean Marketing Review*, 16(4), 71-85.
- Prag, Jay and James Casavant (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. Abraham (1999), "Information, Blockbusters, and Stars: A Study of the Film Industry," *The Journal of Business*, 72(4), 463-492.
- Reddy, Srinivas K, Swaminathan, Vanitha, and Carol M. Motley (1998), "Exploring the Determinants of Broadway Show Success," *Journal of Marketing Research*, 35(3), 370-383.
- Salganik, Matthew J. and Duncan J. Watts (2008), "Leading the Herd Astray: An Experimental Study of Self-fulfilling Prophecies in an Artificial Cultural Market," *Social Psychology Quarterly*, 71(4), 338-355.
- Simonet, T (1978), "Regression analysis of prior experiences of key production personnel as predictor of revenues from high-grossing motion picture in American release," *Amo Press*.
- Sochay, Scott (1994), "Predicting the Performance of Motion Pictures," *Journal of Media Economics*, 7(4), 1-20.
- Wallace, W. Timothy, Alan Seigerman, and Morris B. Holbrook (1993), "The role of actors and actresses in the success of films: how much is a movie star worth?," *Journal of Cultural Economics*, 17(1), 1-27.
- Yoo, Hyunsuk (2002), "The Determinants of Motion

Pictures Box Office Performances - For Movies Produced in Korea Between 1988 and 1999,” Korean Journal of Journalism and Communication Studies, 46(3), 183-213.

Zhao, Min and Jinhong Xie (2011), “Effects of Social and Temporal Distance on Consumers' Responses to Peer Recommendations,” Journal of Marketing Research, 48(3), 486-496.

Zhao, Min, Liu, Nathan and Luheng He (2013), “Social temporal collaborative ranking for context aware movie recommendation,” ACM Transactions on Intelligent Systems and Technology, 4(1).

# The Effects of Consumer Opinion on Box Office Performance in the Korean Movie Market: An Opinion Mining Application

Bumsoo Kim\*, Joo Hwan Seo\*\*

## ABSTRACT

The movie industry is one of the biggest entertainment industries which involve many decisions from various interested parties. However, it also is an extremely risky venture and there has been great interest in identifying the key factors for box office success. There are many possible reasons behind a successful movie such as the participating director and actors, the narrative, the number of opening weeks and so on. With the proliferation of online technology in the 21st century, online word-of-mouth(eWOM) has also been generating a lot of interest in explaining the success of movies. Consequently, a good “buzz” generated in the online community may spread quickly, creating awareness of the movie as well as persuading potential movie-goers. Hence, in this study, we verify the effects of traditional movie related variables and investigate the effects of consumer opinion via eWOM variables on movie box office revenue. The findings from this study enrich the understanding of the factors behind box office success and help the development of a more comprehensive production and marketing strategy for the movie industry participants.

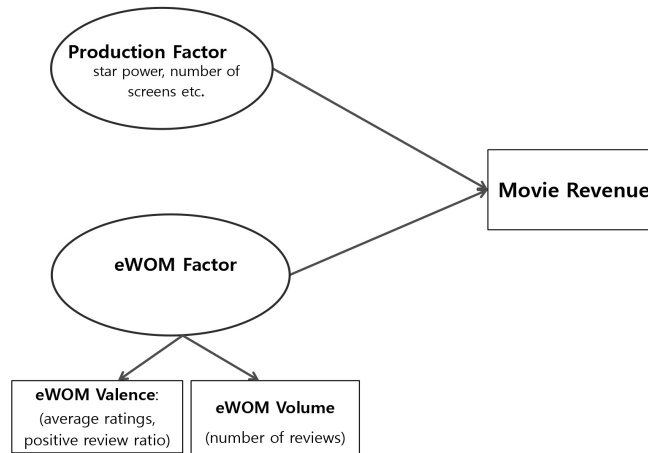
Figure 1 describes our proposed model. We separate the possible factors into two distinct categories, production factor, and eWOM factor, and examine how these factors affect final movie revenue. Production factors are composed of variables which describe the features and distributional characteristics of movies, such as star power, director power, production country, genre, season, film rating and number of opening screens. Increased value of these variables may work as a positive signal to potential movie consumers and in turn positively affect the final movie revenue, In order to analyze the effect of the eWOM factor comprehensively, we include number of reviews, average consumer ratings and ratio of positive consumer reviews as variables. Number of reviews can raise the awareness of the movie whereas average consumer rating and the ratio of positive reviews may have a persuasive effect on potential movie customers. We also investigate if there is a difference between before release eWOM and after release the eWOM effect on final movie revenue. All in all, we hypothesize the following:

---

\* Assistant Professor of Business Science, Sogang University

\*\* Assistant Professor of Marketing, DongA University





<Figure 1> The Proposed Model

H1: eWOM has a positive relationship with box office revenue

H2a: Prerelease and postrelease average movie rating both have positive effect on box office revenue

H2b: Prerelease and postrelease review both have positive effect on box office revenue

H2c: Prerelease and post release positive review ratio have same positive effect on box office revenue

The analyzed data consists of 537 movies released between January 2011 and December 2014. The production factors of these 537 movies are collected from is collected from Korean Film Council database (KOBIS, [www.kobis.or.kr](http://www.kobis.or.kr)), and the eWOM variables are collected from “DAUM Movies” portal site. In addition, we calculate the positive review ratio by the following procedure. First, we create a sentiment dictionary by analyzing the words used in each of 89,401 reviews and differentiate them into three categories; positive, negative and neutral sentiment words. We then, calculate a sentiment score of a review  $j$  consisting of  $N$  meaningful words by utilizing formula (1).

$$Sentiment\ Score_j = \frac{\sum_{i=1}^N Pos.\ Matches_i - \sum_{i=1}^N \neg.\ Matches_i}{\sum_{i=1}^N Pos.\ Matches_i + \sum_{i=1}^N \neg.\ Matches_i} \quad (1)$$

Note that if the sentiment score for review  $j$  is a positive number then we can categorize it as a positive review, whereas a review with a negative sentiment score will be categorized as a negative review. After counting the total positive and negative reviews for a particular movie, one can find the ratio of the positive review ratio. Finally, we perform a multiple linear regression by regressing the production and eWOM variables against log-transformed box office revenue. In addition, we differentiate pre and postrelease eWOM variables and use them separately in order to examine the potential heterogeneity in the eWOM effect. Specifically,

equation (2) and (3) examine prerelease eWOM effects whereas equation (4) and (5) examine the postrelease eWOM effects on box office revenue.

$$\log(sales_i) = \beta_0 + \beta_1 Scr_i + \beta_2 dPow_i + \beta_3 act1Pow_i + \beta_4 act2Pow_i + \beta_5 distA_i + \beta_6 Season_i + \beta_7 Action_i + \beta_8 Comedy_i + \beta_9 DP_i + \beta_{10} DP18_i + \beta_{11} national_i + \beta_{12} AvgbefReview_i + \beta_{13} befReviewVol_i + \epsilon_i \quad (2)$$

$$\log(sales_i) = \beta_0 + \beta_1 Scr_i + \beta_2 dPow_i + \beta_3 act1Pow_i + \beta_4 act2Pow_i + \beta_5 distA_i + \beta_6 Season_i + \beta_7 Action_i + \beta_8 Comedy_i + \beta_9 DP_i + \beta_{10} DP18_i + \beta_{11} national_i + \beta_{12} AvgbefReview_i + \beta_{13} befReviewVol_i + \beta_{14} befReviewPosRatio_i + \epsilon_i \quad (3)$$

$$\log(sales_i) = \beta_0 + \beta_1 Scr_i + \beta_2 dPow_i + \beta_3 act1Pow_i + \beta_4 act2Pow_i + \beta_5 distA_i + \beta_6 Season_i + \beta_7 Action_i + \beta_8 Comedy_i + \beta_9 DP_i + \beta_{10} DP18_i + \beta_{11} national_i + \beta_{12} AvgAftReview_i + \beta_{13} AftReviewVol_i + \epsilon_i \quad (4)$$

$$\log(sales_i) = \beta_0 + \beta_1 Scr_i + \beta_2 dPow_i + \beta_3 act1Pow_i + \beta_4 act2Pow_i + \beta_5 distA_i + \beta_6 Season_i + \beta_7 Action_i + \beta_8 Comedy_i + \beta_9 DP_i + \beta_{10} DP18_i + \beta_{11} national_i + \beta_{12} AvgbefReview_i + \beta_{13} befReviewVol_i + \beta_{14} AftReviewPosRatio_i + \epsilon_i \quad (5)$$

Parameter estimates for each model are summarized in <Table 1>. First of all, one can see that all production factor variables except the director power variable are significantly related to box office revenue. What is also interesting is that the coefficient for the production country is negative, meaning that on average the box office revenue for international movies is lower than that of Korean movies. This result shows that in the Korean movie market, Korean movies have a competitive edge on international movies including Hollywood movies. This is a contradicting result from previous findings of Waterman and Jayakar(2000), where they show that in many countries Hollywood movies perform well compared to the respective domestic movies due to the vast difference in production cost.

In regards to the effect of eWOM, we find that other than prerelease average review point all eWOM variables are positively related to box office revenue regardless of whether they are prerelease reviews or postrelease reviews. Thus we conclude that hypothesis 1 is supported. When we look at the detail components of eWOM effects, we first see that review volume is positively related to box office revenue in all models. This result confirms our belief that the number of reviews raise consumer awareness and in turn positively affect box office revenues. Thus we conclude that hypotheses 2b) are supported. There are some interesting findings regarding the average review point variable. We see that prerelease average point variable has no effect on box office revenues whereas postrelease average point variable has a significant positive relationship with box office revenues. This difference can be explained in terms of reliability of the average review point measure. Prerelease average review point is based on the indirect experience of the movie, such as news articles and previews, and therefore potential consumers are not as persuaded by prerelease average review point in making their decisions.

&lt;Table 1&gt; Model Parameter Estimates

	Model 1		Model 2		Model 3		Model 4	
	Coefficient	VIF	Coefficient	VIF	Coefficient	VIF	Coefficient	VIF
Intercept	19.758** (0.492)		21.190** (0.669)		18.596** (0.477)		20.416** (0.673)	
# of Screens	0.0034** (0.0006)	1.112	0.0033** (0.0006)	1.112	0.0033** (0.0006)	1.114	0.0032** (0.0006)	1.114
Director Power	0.0026 (0.011)	1.156	0.0021 (0.011)	1.157	0.0062 (0.0109)	1.148	0.0046 (0.0108)	1.150
Actor1 Power	0.0190* (0.0095)	1.184	0.0179* (0.0089)	1.186	0.0198* (0.0093)	1.182	0.0194* (0.0092)	1.182
Actor2 Power	0.0174* (0.0084)	1.219	0.0197* (0.0083)	1.229	0.0148* (0.007)	1.219	0.0174* (0.0082)	1.229
A level Distributor	1.062** (0.161)	1.086	1.091** (0.159)	1.090	0.994** (0.157)	1.087	1.101** (0.1557)	1.087
Season	0.304* (0.138)	1.028	0.292* (0.137)	1.029	0.292* (0.135)	1.029	0.279* (0.133)	1.029
Action	0.459** (0.167)	1.329	0.423* (0.167)	1.336	0.521** (0.165)	1.329	0.486** (0.163)	1.334
Comedy	-0.346 (0.198)	1.171	-0.288 (0.197)	1.182	-0.307 (0.195)	1.173	-0.224 (0.193)	1.188
G rated	-0.606** (0.188)	1.414	-0.566** (0.187)	1.420	-0.753** (0.186)	1.433	-0.724** (0.184)	1.435
R rated	-0.910** (0.177)	1.131	-0.916** (0.176)	1.131	-0.823** (0.175)	1.147	-0.822** (0.174)	1.147
International	-0.971** (0.169)	1.352	-0.998** (0.167)	1.356	-1.013** (0.165)	1.352	-1.035** (0.164)	1.353
Average Review Point(Before Opening)	0.079 (0.087)	1.047	0.1021 (0.089)	1.204				
Average Review Point(After Opening)					0.2651** (0.054)	1.127	0.3115** (0.0549)	1.186
Review Volume(Before Opening)	0.0098** (0.0027)	1.075	0.0093** (0.0027)	1.078				
Review Volume(After Opening)					0.0004** (0.0001)	1.089	0.0004** (0.0001)	1.093
Positive Review Ratio(Before Opening)			2.187** (0.699)	1.109				
Positive Review Ratio(After Opening)							2.471** (0.653)	1.111
F-test statistic	23.52**		22.91**		25.6**		25.4**	
AdjR <sup>2</sup>	0.343		0.365		0.373		0.389	

\* :  $p < 0.05$ , \*\* :  $p < 0.01$  ;

In comparison, postrelease average review point is mostly based on actual movie experiences and therefore it can be argued that postrelease average review point is a more accurate representation of the true satisfaction of customers. In sum, we find that hypothesis 2a) is not supported. Finally, we find that the coefficient for positive review ratio is significant in both prerelease and postrelease scenarios. This shows that positive review ratio is

more effective in measuring the valence of the eWOM effect than average review point and confirms the belief that positive eWOM affect box office revenues. Therefore, we conclude that hypothesis 2c) is supported from our results.

In this study, we examined the effects of consumer opinion via eWOM factors on movie box office revenues while controlling for production related effects. We find that the production factor related variables and eWOM factor variables are significant in explaining the differences in the success of individual movies. Furthermore, we find that the prerelease review volume and the prerelease review ratio have a positive relationship with box office revenue. Movie industry managers can utilize these findings in focusing their marketing efforts more efficiently and in devising a coordinated marketing strategy.

Keywords : Movie Success, Data Mining, eWOM, Online Review, Consumer opinion analysis