

Mixed-effects Bass 모형을 이용한 영화확산모형: 온라인구전의 영향을 중심으로

주영진*

본 연구에서는 영화진흥위원회 통합전산망의 영화박스오피스자료와 온라인영화사이트의 리뷰자료를 이용하여 온라인구전을 중심으로 주요 영향요인들이 영화확산에 미치는 영향을 분석하였다.

온라인구전이 영화성과에 미치는 영향을 분석한 많은 선행연구들에서는 온라인구전의 주요 차원인 온라인구전의 양과 방향성의 영향력을 비교하고 있으나, 분석결과는 분석에 사용된 자료나 모형에 따라 다르게 제시되고 있다. 2차자료는 연구자가 사전에 정한 연구설계에 따라 변수의 측정단위가 잘 통제될 수 있는 1차자료와 달리 다양한 패널시계열특성들을 지니기에, 2차자료를 이용한 연구에서는 분석모형이 사용되는 자료의 패널시계열특성을 효과적으로 처리할 수 있어야 한다. 그리고, 2차자료를 이용하여 영화성과를 분석한 연구들에서는 일반적으로 높은 왜도를 지니는 영화성과를 조정하기 위해 로그변환한 영화성과를 선형모형으로 분석하고 있다. 매출이나 관람객수가 매우 큰 흥행실적 상위 영화들의 중요성을 분석과정에서 고려하기 위해서는 로그변환을 하지 않은 영화성과 자체를 비선형모형으로 분석할 필요가 있다. 선행연구들 중에서는 비선형모형으로 분류될 수 있는 Bass 모형을 적용한 연구들이 일부 있으나, 상대적으로 많은 연구들에서는 선형모형인 회귀모형이 적용되어 왔다.

이에 본 연구에서는 한국영화시장에서 2018년부터 2021년까지 4년간 개봉된 연도별 박스오피스 상위 30편에 해당하는 총 118편의 영화를 대상으로 한 주별패널자료에 Mixed-effects Bass 모형을 적용하여 로그변환을 하지 않은 영화성과(관람객수) 자체가 나타내는 영화확산에 개봉첫주(더미변수), 스크린수, 리뷰수, 평점 등의 영향요인들이 미치는 영향을 분석하였다. Mixed-effects Bass 모형은 분석에 사용된 영화패널자료에 내재된 주된 시계열적 특성인 장기적 확산과정은 Bass 모형이 내재하고 있는 확산과정을 통해 반영하고, Bass 모형의 모수인 혁신계수와 모방계수가 영화패널별로 고정효과인 스크린수, 리뷰수, 평점 등의 영향요인들과 변동효과인 오차항에 의해 설명될 수 있는 계층적비선형모형이다. 분석결과 영화확산과정의 혁신계수는 개봉첫주와 스크린수에 유의한 정(+)의 영향을 받고, 모방계수는 리뷰수와 평점에 유의한 정(+)의 영향을 받는 것으로 나타났다.

본 연구의 주요 의의는 다음과 같이 요약될 수 있다. 먼저 본 연구는 온라인구전 등이 영화성과(관람객수)에 미치는 영향을 주별패널자료를 이용하여 분석함에 있어 확산과정을 나타내는 Bass 모형에 대한 추정과 여러 영향요인들과 오차항이 Bass 모형의 모수인 혁신계수와 모방계수에 미치는 영향을 나타내는 회귀식에 대한 추정을 통합하여 추정된 Mixed-effects Bass 모형을 처음으로 적용한 연구이다. 그리고 본 연구는 분석결과 온라인구전의 양(리뷰수)과 방향성(평점)이 모두 통계적으로 유의한 영향력을 지니는 것을 확인하였다.

주제어 : 온라인구전, 영화확산, Mixed-effects Bass 모형, 계층적비선형모형, 영화패널자료

I. 서론

성공적인 온라인마케팅전략을 위해 온라인구전(eWOM: electronic(or online) word of mouth)의

중요성이 증가하며 온라인구전의 효과에 대한 연구들이 활발히 수행되어 왔다. 전통적인 구전에 대한 연구가 주로 실험설계 등을 통한 1차자료를 이용한 연구가 많은 반면, 온라인구전에 대한 연

* 충북대학교 경영학부 교수(yjjoo@cbnu.ac.kr)

구는 실제 온라인사이트에서 수집한 2차자료를 이용한 연구가 많은 편이다. 2차자료를 이용한 온라인구전에 대한 연구가 많아진 배경에는 온라인 상에서 기존 소비자들의 소비경험을 편리하게 소통할 수 있는 다양한 사이트들이 등장하고, 이들 사이트에 모여 있는 구전정보들을 쉽게 수집하기 위한 정보기술이 급격히 발전 및 보급되고 있기 때문으로 볼 수 있다.

구전에 대한 연구에서 중요하게 취급되어 온 내용의 하나로 구전의 차원을 꼽을 수 있다(류강석, 유시진, 2015). 구전의 차원은 구전이 얼마나 많거나 적은지를 구분하는 구전의 양(volume)과 구전의 내용이 긍정적인지 부정적인지를 구분하는 구전의 방향성(또는 극성) (valence) 두 가지가 가장 대표적이며, 구전의 분산(variance)이나 구전의 감성(sentiment) 등의 차원들이 추가되기도 한다. 구전의 양과 방향성 중에서 Duan et al.(2008a, 2008b), Liu(2006) 등은 구전의 양을 중시하고 있지만, Chintagunta et al.(2010) 등은 구전의 방향성을 중시하는 것과 같이 구전의 차원별 영향력은 선행연구들에서 일관되게 나타나지 않고 있다.

구전의 차원별 영향력의 차이는 이론적 맥락의 차이에서 발생할 수도 있지만, 분석모형이나 분석에 사용된 자료의 차이에서 발생할 수도 있다(You et al., 2015). 주영진과 최유리(2023)는 2차자료를 이용한 온라인구전 연구들에서 구전의 차원별 영향력의 차이가 나타나는 이유로 구전의 차원으로 조작된 변수의 측정단위가 다른 것을 분석모형이 적절하게 고려하고 있지 못하기 때문으로 설명하였다. 실험설계를 통해 1차자료를 직접 수집한 연구에서는 구전의 차원들이 연구설계에 맞춰 서로 비교 가능한 단위로 통제되어 측정될 수 있으나, 2차자료를 이용한 연구에서는 구전의 차원들로 조작된 변수들의 패널시계열적 특성이 다를 수 있어 이러한 차이를 고려하지 못할 경

우 분석결과가 왜곡될 수 있다는 것이다.

한편, 2차자료를 이용한 구전에 대한 연구들 중 많은 연구들에서는 성과변수의 왜도(skewness)가 높은 점을 고려하여 성과변수를 로그변환(log transformation)한 값을 대상으로 선형모형을 이용하여 분석하는 것이 일반적이다. 로그변환을 하지 않은 성과변수 자체를 분석한 연구들에는 차경천 외(2014), Dellarocas et al.(2007) 등과 같이 비선형모형으로 분류될 수 있는 Bass 모형(Bass, 1969)을 이용한 연구들이 일부 있으나, 로그변환한 성과변수를 선형모형으로 분석한 연구들에 비해 크게 부족한 편이다. 구전의 차원별 영향력이 선행 연구들에서 일관되지 못한 상황에서 성과변수 자체를 비선형모형을 이용하여 분석하기 위한 연구가 더 많이 진행될 필요가 있다. 로그변환한 성과변수를 분석하는 것은 상대적으로 매우 큰 일부 극단적인 관찰치들이 모형의 추정에 민감하게 작용하는 것을 완화할 수 있는 장점이 있지만, 성과변수가 매우 큰 일부 극단적 관찰치가 오히려 중요하게 설명되어야 할 관찰치라는 점을 주의할 필요가 있다.

이에 본 연구에서는 한국의 영화시장에서 수집한 2차자료를 이용하여 영화성과인 로그변환을 하지 않은 영화관람객수 자체를 설명하기 위한 비선형모형을 적용하여 온라인구전의 차원별 영향력을 분석하고자 하였다. 이를 위해 본 연구에서는 한국의 영화시장에서 2018년부터 2021년까지 4년간 각 연도별 박스오피스 상위 30편에 속하는 총 118편의 영화들을 대상으로 수집된 영화관람객수, 스크린수, 온라인구전의 양(리뷰수)과 방향성(평점) 등에 대한 주별패널자료를 Mixed-effects Bass 모형을 이용하여 분석하였다. 본 연구에서 적용한 Mixed-effects Bass 모형은 비선형모형인 Bass 모형의 모수가 스크린수, 리뷰수, 평점 등의 영향요인에 의한 고정효과(FE: fixed effects)와 오

차에 의한 변동효과(RE: random effects)가 결합된 혼합효과(mixed-effects)를 지니는 것을 나타낸 모형으로 Bass 모형식과 모수에 대한 회귀식이 단계별로 이루어진 계층적비선형모형(hierarchical nonlinear model)이다.

본 연구는 2장에서 영화확산에 대한 온라인구전의 영향과 영화확산에 대한 확산모형 적용 등을 중심으로 선행연구를 살펴보고, 본 연구의 연구문제를 제시하였다. 3장에서는 본 연구에서 사용한 영화패널자료와 본 연구에서 적용한 Mixed-effects Bass 모형 등 연구방법을 설명하고 있다. 4장은 Mixed-effects Bass 모형을 적용한 영화패널자료에 대한 분석결과와 해석을 제시하였고, 마지막 5장은 본 연구의 결론, 시사점과 한계점 등을 제시하고 있다.

II. 연구배경 및 연구문제

1. 영화확산과 온라인구전

최근 영화산업을 대상으로 실제 자료를 이용한 실증연구들이 크게 증가하고 있는데, 이는 웹스크래핑을 이용해서 Box Office Mojo, IMDb 등의 온라인 영화사이트에서 영화관련 2차자료들을 쉽게 수집할 수 있기 때문이기도 하다(McKenzie, 2022). 이에 따라 이전에 주로 설문이나 실험설계를 통해 진행되었던 온라인구전이 영화확산에 미치는 영향에 대한 연구에서도 2차자료를 이용하는 비중이 증가하고 있다(McKenzie, 2022). 온라인구전이 영화확산에 미치는 영향을 분석한 연구들은 배정호 외(2010), 이중원과 박철(2014), 주영진과 최유리(2021, 2023), 차경천 외(2014), Chintagunta et al.(2010), Dellarocas et al.(2007), Duan et al.(2008a, 2008b), Liu(2006) 등을 꼽을 수 있다.

영화확산에 대한 온라인구전의 영향을 다룬 연구들에서는 온라인구전의 두가지 핵심적 차원인 구전의 양(volume)과 구전의 방향성(valence)의 영향력에 대한 차이가 중요하게 분석되어 왔다(류강석, 유시진, 2015; You et al., 2015). Liu(2006)는 영화소비자의 인지행동경로(cognition-behavior routes)에서 구전의 양(volume)은 정보적(informative) 역할을 통해 소비자의 인지(awareness)를 강화하고, 구전의 방향성(valence)은 기본적으로 긍정 또는 부정의 방향성을 지님으로써 소비자의 태도에 영향을 주는 설득적(persuasive) 역할을 하는 것으로 구분하고 있다.

기존 연구들에서 온라인구전의 양과 방향성의 영향력은 일관되게 나타나지 않는다. Duan et al.(2008a, 2008b), Kim et al.(2019), Liu(2006) 등의 연구는 영화매출에 대해 구전의 양(volume)은 유의한 영향을 미치지만 구전의 방향성(valence)은 유의한 영향을 미치지 않는다고 했지만, 배정호 외(2010), Chintagunta et al.(2010) 등의 연구는 영화매출에 대해 구전의 방향성(valence)은 유의한 영향을 미치지만 구전의 양(volume)은 유의한 영향을 미치지 않는다는 상반된 결과를 제시하고 있다. 구전의 양과 방향성에 대한 상반된 실증 분석 결과에 대해 Chintagunta et al.(2010)은 지역별로 서로 다른 영화 개봉시기를 반영하지 못한 합산 오류(aggregation bias) 때문으로, 배정호 외(2010)는 온라인구전과 영화매출 간의 양방향적 인과관계 때문으로 설명하고 있다. 주영진과 최유리(2023)는 영화패널자료에 대한 총량회귀모형과 주별패널분석모형의 분석결과를 비교하여 영화패널자료에 내재된 패널간 이질성(heterogeneity)과 시계열적 특성이 효과적으로 고려되는지의 여부에 따라 구전의 양과 방향성 등의 온라인구전의 차원별 영향력이 다르게 나타날 수 있음을 보이고 있다. 즉, 2차자료를 이용한 온라인구전이 영

화확산에 미치는 영향에 대한 기존연구들에서 구체적인 자료특성이나 분석모형에 따라 온라인구전의 양(volume)과 방향성(valence)의 영향력이 다르게 분석될 수 있는 것이다. You et al.(2015)의 메타 연구에서도 분석모형과 자료의 특성이 온라인구전의 양(volume)과 방향성(valence)의 영향력을 결정하는 주요 요인의 하나로 분석되었다.

한편, 2차자료를 이용한 온라인구전이 영화확산에 미치는 영향에 대한 기존연구들에서 사용된 분석모형은 회귀분석(regression analysis)을 중심으로 한 선형모형과 확산모형(diffusion model)을 중심으로 한 비선형모형으로 나뉜다. 영화매출이나 영화관람객수 등 영화성과를 나타내는 변수들은 블록버스터급 영화와 다수의 일반 영화의 차이, 개봉 후 시간의 변화에 따른 차이 등으로 인해 높은 왜도(skewness)를 지닌다. 이에 따라 많은 연구들에서는 영화성과에 로그변환(log transformation)을 적용하여 높은 왜도를 조정하여 분석하였다(McKenzie, 2022). 로그변환한 영화성과에 대한 분석은 선형모형을 통한 분석이 적합하지만, 상대적으로 높은 성과의 표본에 대한 설명력이 실제보다 과소평가될 수 있는 위험이 있다. 로그변환을 적용하지 않은 영화성과 자체를 대상으로 한 분석은 개봉시점부터 시간이 지남에 따라 비선형적으로 변화하는 확산과정을 설명할 수 있어야 하는데, 이를 위해 차경천 외(2014), Dellarocas et al.(2007) 등의 연구에서와 같이 Bass 모형(Bass, 1969)을 확장한 모형이 적용되어 왔다.

2. 확산모형과 영화확산

확산모형(diffusion model)은 혁신신제품이 시장에 도입되어 성장하고 성숙기를 거쳐 쇠퇴하는 일련의 확산과정(diffusion process) 또는 제품수명주기(product life cycle)을 잘 추종할 수 있는 비선

형함수를 이용하여 혁신신제품의 확산과정을 설명하기 위한 모형이다. 마케팅분야에서 가장 널리 이용되고 있는 확산모형은 Bass 모형(Bass, 1969)을 꼽을 수 있다. Bass 모형(Bass, 1969)은 특정 혁신신제품의 잠재시장에 속한 특정소비자가 특정시점에서 해당 혁신신제품을 그 시점까지는 구매하지 않았는데 그 시점에 구매할 조건부확률(또는 hazard 함수)을 혁신효과(innovation effect)와 모방효과(imitation effect)의 합으로 정의함으로써 해당 혁신신제품의 확산과정을 모형화하였다. Bass 모형에 대한 상세한 이해는 Mahajan et al.(1990)를 참고할 수 있으며, Bass 모형은 Mahajan and Muller(1979), Mahajan et al.(1990) 등에서 고찰하고 있는 것과 같이 수십년동안 마케팅분야의 다양한 혁신신제품의 확산과정을 설명하는 데 중요하게 적용되어 왔다.

확산모형은 김태구와 홍정식(2015), 이유석 외(2017), 장광필(2008), 차경천 외(2014), Dellarocas et al.(2007), Lee et al.(2022) 등의 연구에서와 같이 영화의 확산을 설명하기 위한 목적으로도 효과적으로 적용되어 왔다.

Dellarocas et al.(2007)은 Yahoo! Movies, Box Office Mojo, Hollywood Reporter 등에서 수집한 영화박스오피스 자료와 관람객 및 비평가의 리뷰 자료를 Bass 모형을 기반으로 개발한 단계별모형을 MCMC(Markov Chain Monte Carlo) 방법으로 추정하여 분석하였다. Dellarocas et al.(2007)의 단계별모형은 기본적인 영화확산과정을 Bass 모형에서 혁신계수(p)가 도입 이후 경과한 시간에 따른 감소하는 경향을 반영하기 위한 할인요소(discount factor)를 결합한 형태($p\delta^t$)로 제안하고, 제안된 모형의 모수들인 잠재시장규모, 혁신계수, 모방계수 및 할인요소 등을 여러 영향 요인변수들의 선형함수인 회귀식으로 정의하였다. 이를 통해 Dellarocas et al.(2007)은 영화에 대한 초기의 온라인 리뷰

양(volume)이 초기 영화매출에 대한 대리변수(proxy)로 유용하고, 온라인 리뷰의 평균평점(average valence)이 모방효과에 유의한 정(+)의 효과가 있다는 등의 결과를 제시하고 있다.

장광필(2008)은 2003년 11월부터 2004년 11월까지 한국에서 개봉된 영화들 중 41편의 영화를 대상으로 주별관객수의 확산과정을 Bass 모형으로 추정하고, 각각의 영화들에 대해 별도로 추정된 확산모형을 바탕으로 계산된 혁신수용자와 모방수용자를 영화에 대한 기대, 개봉 전 인지도, 영화 품질, 기대 불일치 등의 변수로 설명하기 위한 회귀분석 실시하였다. 이를 통해 장광필(2008)은 개봉전 인지도와 높은 기대수준이 영화의 혁신수용자를 증가시키고, 영화에 대한 기대와 성과의 불일치에서 오는 구전효과가 모방수용자를 증가시키고 있음을 실증하였다. 다만, 장광필(2008)은 Bass 모형으로 추정한 영화의 확산과정에 여러 영향요인들이 미치는 영향을 Bass 모형의 모수를 통해 결합하는 대신 Bass 모형의 추정식을 바탕으로 계산된 혁신수용자의 수와 모방수용자의 수를 여러 영향요인들로 설명하는 방법을 적용함으로써 단계별 모형을 통합하여 정의하고 추정하기 어려운 한계를 지닌다.

차경천 외(2014)는 영화진흥위원회와 네이버에서 수집한 자료를 이용하여 2012년 한국에서 개봉한 13편의 영화를 대상으로 일별 영화관람객의 확산 과정에 대한 온라인구전의 효과를 GBM(generalized Bass model) (Bass et al., 1994)의 형태로 개발하여 분석하였다. 차경천 외(2014)는 Bass 모형의 모방계수에 온라인구전(댓글)의 양, 평점, 동의정도 등을 Cobb-Dougllass 함수꼴로 결합한 구전자산항을 더한 형태의 모형에 주말과 공휴일 효과를 GBM처럼 추가한 모형을 제안하고, 이를 13편의 영화에 대해 개별적으로 추정한 결과를 비교하고 있다. 이를 통해 차경천 외(2014)는 영화확산에

대한 구전자산 개념의 유용성을 실증적으로 제시하였으나, 분석대상 영화들에 대해 개별적으로 추정한 결과를 비교함에 따라 분석결과의 일반화가 어려울 수 있는 한계를 지닌다.

이유석 외(2017)는 2014년과 2015년 한국에서 개봉한 누적 관객수 상위 50편의 영화를 대상으로 일별 영화관람객의 확산과정을 GBM(generalized Bass model) (Bass et al., 1994)의 형태로 추정하고, 추정된 Bass 모형의 모수들인 혁신계수, 모방계수와 잠재시장규모에 스크린수, 시사회 관객수, 개봉 전 기대지수, 배급사규모 등이 미치는 영향을 SUR(seemingly unrelated regression) 모형을 이용하여 분석하였다. 이를 통해 이유석 외(2017)는 개봉 전 시사회의 영향력과 개봉 시 스크린 확보의 중요성 등을 실증적으로 검증하였다. Lee et al.(2022)은 한국영화시장을 대상으로 2005년부터 2019년까지 매년 상위 20편의 한국영화와 상위 20편의 외국영화 등 모두 600편의 영화를 대상으로 이유석 외(2017)의 연구에서 사용된 것과 유사한 모형을 이용하여 15년간의 시간적 변화에 따른 영화확산과정에서의 확산계수와 모방계수의 변화를 실증적으로 분석하였다. 다만, 이유석 외(2017)와 Lee et al.(2022)은 GBM에 대한 추정과 Bass 모형 모수에 대한 SUR모형에 대한 추정을 단계별로 나누어서 추정하고 있다는 점에서 단계별모형을 MCMC(Markov Chain Monte Carlo) 방법으로 통합하여 추정한 Dellarocas et al.(2007)과 구분되고 있다.

김태구와 홍정식(2015)은 2011년 9월 22일부터 2013년 2월 3일까지 한국에서 개봉한 175편의 영화와 미국에서 개봉한 123편의 영화를 대상으로 한국과 미국의 영화확산과정을 비교 분석하였다. 특히, 김태구와 홍정식(2015)은 영화확산과정을 설명하기 위한 모형으로 Bass 모형에 적분상수를 추가한 BMIC(Bass model with integration constant)

모형을 이용하였는데, 이는 영화의 확산과정이 전형적인 확산과정과 달리 확산초기부터 수용자의 수가 매우 큰 것을 반영하기 위한 것이다. 다만, 김태구와 홍정식(2015)은 확산모형의 모수에 대한 한국과 미국에서의 차이를 비교 분석함으로써 두 시장에서 확산패턴의 차이가 있음을 보이고는 있지만, 그런 차이가 어떤 영향요인들에 의해 발생한 것인지에 대한 탐색은 제시되고 있지 않는다.

한편, 이유석 외(2017), 차경천 외(2014), Dellarocas et al.(2007), Lee et al.(2022) 등에서는 확산모형의 모수들에 대한 여러 영향요인들의 영향력을 단계별모형을 이용하여 분석하고 있다. 이와 유사하게 Bass 모형(Bass, 1969)을 확장한 연구들 중 기업의 마케팅전략 실행 결과와 같은 여러 영향요인들이 혁신신제품의 확산과정에 미치는 영향을 분석하기 위한 목적으로 수행되었던 대표적인 연구들로는 Bass et al.(1994), Horsky and Simon(1983), Joo and Jun(1996) 등을 꼽을 수 있다. Horsky and Simon(1983)은 광고가 확산과정을 촉진시키는 것을 보이고 있으며, Bass et al.(1994)은 Bass 모형의 hazard 함수에 여러 마케팅 변수의 영향을 결합하기 위한 GBM(generalized Bass Model)을 제시하였다. Joo and Jun(1996)의 연구에서도 Bass 모형의 확산과정이 단기적인 시장환경의 변화에 따라 일시적으로 촉진되거나 둔화될 수 있음을 한국시장의 가정용에어컨의 확산을 이용하여 실증적으로 보이고 있다.

3. 연구문제

영화확산에 대한 온라인구전의 두가지 핵심적 차원인 구전의 양(volume)과 방향성(valence)은 각각 정보적 역할과 설득적 역할을 통해 소비자가 영화를 알고 긍·부정의 태도를 형성하는 데 영향을 미치는 것으로 기대되어 왔다(Liu, 2006). 이

에 따라 많은 연구들에서 영화박스오피스 및 온라인 영화사이트에서 수집한 2차자료를 이용하여 온라인구전의 양(volume)과 방향성(valence)의 영향력에 대한 실증연구들이 수행되었다.

영화를 대상으로 한 온라인구전을 다룬 기존 연구들 중 일부 연구들(Duan et al., 2008a, 2008b; Kim et al., 2019; Liu, 2006)에서는 구전의 양(volume)이, 다른 일부 연구들(배정호 외, 2010; Chintagunta et al., 2010)에서는 구전의 방향성(valence)이 더 중요한 것으로 분석되었다. You et al.(2015)은 온라인구전의 양(volume)과 방향성(valence)의 영향력의 차이가 분석에 사용된 자료나 모형의 특성에 따라서도 발생된다는 것을 제기하였다. 주영진과 최유리(2023)는 2차자료를 이용한 영화확산에 대한 온라인구전의 차원별 영향력을 비교한 분석을 통해 영화패널자료의 패널시계열적 특성을 효과적으로 내재한 모형이 적용될 필요가 있음을 보였다. 주영진과 최유리(2023)는 주별패널자료를 이용하여 영화관람객수(로그값)가 스크린수(로그값)와 리뷰수(로그값)와 공적분(co-integration) 관계가 있음을 반영하기 위한 패널오차수정모형(panel error correction model)을 적용한 분석을 통해 온라인구전의 차원들 중 리뷰수로 측정된 구전의 양(volume)만이 통계적으로 유의하게 영화관람객수에 정(+)의 영향을 미친다고 하였다. 그러나 주영진과 최유리(2023)에서 적용된 방법은 시계열 관측치가 충분하지 않은 경우 적용이 어렵고, 영화관람객수의 로그값을 설명하는 것이 추정과정에서 상대적으로 관람객수가 많은 박스오피스 상위의 영화나 모든 영화들에 대해서 개봉초기의 관측치에 대한 가중치를 감소시킬 수 있는 한계를 지니고 있다.

영화관람객수(또는 영화매출액)의 로그값을 대상으로 한 선형모형 대신 영화관람객수(또는 영화매출액) 자체를 대상으로 한 비선형모형에 의한

분석이 필요하다. 영화관람객수(또는 영화매출액) 자체는 전형적으로 영화가 개봉된 초기에 관람객수가 많았다가 점차 시간이 지나면서 관람객수가 급격하게 감소하는 모습을 보이는데, 이러한 비선형적 시계열적 특성은 확산모형(diffusion model)을 이용하여 설명할 수 있다. 차경천 외(2014), Dellarocas et al.(2007) 등이 영화를 대상으로 온라인구전을 분석하기 위해 확산모형을 적용하였으나, 상대적으로 선형모형에 의한 분석에 비해 확산모형 등 비선형모형에 의한 분석이 크게 부족한 편이다.

이에 본 연구에서는 영화확산에 대한 온라인구전의 영향력을 분석하기 위해 영화패널자료의 비선형적 시계열적 특성을 반영한 분석모형을 탐색하기 위한 목적으로 다음의 연구문제 1을 제시하였다.

연구문제 1: 영화확산과정에서 온라인구전의 영향력을 분석하기 위한 확산모형(diffusion model)을 탐색·제안한다.

이어서 연구문제1을 통해 제안된 확산모형(diffusion model)을 이용하여 온라인구전의 두가지 핵심적 차원인 구전의 양(volume)과 방향성(valence) 등의 영향요인들이 영화확산에 미치는 영향을 분석할 필요가 있다. Bass 모형에 기반한 단계별 모형을 통해 Dellarocas et al.(2007)은 온라인구전의 양(리뷰수)과 방향성(평균평점)이 Bass 모형의 혁신계수와 모방계수에 통계적으로 유의한 정(+)의 영향을 미치는 것으로 분석하였고, 차경천 외(2014)는 온라인구전자산을 구성하는 댓글수, 평점, 댓글동의 정도 등을 Bass 모형의 모방계수에 결합한 GBM (generalized Bass model)을 적용한 분석 결과를 제시하고 있으나 온라인구전의 개별 요소들의 영향력은 영화별로 따로 추정하여 일반적인 해석이 제한적인 한계를 지니고 있다. 영화확산에 대한 온라인구전의 영향을 분석한 많은 연구들이 선형모형에 편중된 상황에서 비선형모형인 확산

모형을 적용하여 온라인구전의 영향을 분석하기 위한 연구가 추가될 필요가 있다.

이에 본 연구에서는 위의 연구문제1을 통해 제안된 확산모형(diffusion model)을 기반으로 영화확산에 대한 온라인구전의 영향력을 온라인구전의 양(volume)과 방향성(valence)을 중심으로 분석하기 위한 목적으로 다음의 연구문제 2를 제시하였다.

연구문제 2: 확산모형(diffusion model)을 기반으로 영화확산과정에서 온라인구전 양(volume)과 방향성(valence) 등의 영향력을 분석한다.

III. 연구방법

1. 자료

분석에 사용된 자료들 중 각 영화들의 관람객수와 스크린수는 영화진흥위원회 통합전산망(<https://www.kobis.or.kr>)에서 제공하는 일별자료를 사용하였고, 온라인구전에 대한 리뷰수와 평점은 다음 영화사이트(<https://movie.daum.net>)에서 스크래핑한 자료를 사용하였다. 두 사이트에서 수집한 영화별 관람객수, 스크린수, 리뷰수, 평점 등의 자료는 월요일부터 일요일까지를 기준으로 한 주별패널자료로 합산되었다. 즉, 영화별 관람객수와 스크린수는 해당 주별로 일별자료를 합산한 것이고, 리뷰수와 평점은 리뷰와 평점이 입력된 일/시/분/초 단위의 시간정보를 이용하여 해당 주별로 합산하거나(리뷰수) 해당 주별로 평균한(평점) 것이다. 개봉되는 영화들의 개봉 요일이 모두 월요일은 아니고, 오히려 수요일 정도가 많은 것으로 나타나지만, 영화 관람객수의 일별 시계열이 7일 주기의 강한 계절성(seasonality)을 지니며 주로 금요일부터 일요일까지의 주 후반에 높은 수를 나타내는 점을 반영하여 일요일까지의 한주를 기준으

로 정한 것이다. 이유석 외(2017), 차경천 외(2014), Lee et al.(2022) 등은 일별시계열을 분석하며 관람객수가 많은 주말에 대한 더미변수를 포함하여 일별시계열에 포함된 7일주기의 계절성을 제거하였다. 그러나 일반적으로 영화 관람객수에 대한 일별시계열은 개봉 첫주 주말에 확산과정의 정점(peak)을 형성하게 되는데, 그러한 정점이 개봉 후 관람객수가 증가하다 생긴 정점인지 또는 높은 대기수요에 따라 개봉과 함께 나타난 정점이 점진적으로 감소하는 확산과정에서 요일에 의한 계절성으로 부풀려진 것인지를 모형이 구분하는 것이 한계가 있을 수 있다. 이에 본 연구에서는 요일에 의한 7일주기의 계절성을 분석에서 배제하기 위한 또 다른 방법으로 일별시계열 대신 주별시계열을 사용한 것이다.

한편, 위의 2.1절에서 고찰한 것과 같이 영화를 대상으로 수행된 많은 기존 연구들에서는 매출액이나 영화관람객수와 같은 영화성과변수들의 높은 왜도(skewness)를 조정하고자 로그변환(log transformation)한 영화성과를 분석하였다. 그러나 본 연구에서는 로그변환을 하지 않은 영화성과의 원래 값이 확산 과정에 따라 크게 차이가 나는 점을 비선형모형인 확산모형을 이용하여 분석하고자 하였다.

2. 모형

혁신신제품의 확산과정을 설명하기 위한 연구들 중 가장 영향력이 큰 연구인 Bass(1969)는 시장의 대표 소비자가 특정 시점(t)에서 그때까지 해당 혁신신제품을 수용하지 않고 있다 바로 그 시점에서 수용하는 조건부 확률인 hazard 함수 $f_t/(1-F_t)$ 를 혁신효과와 모방효과의 합($p+qF_t$)으로 정의하고, 이를 동질적 특성의 잠재소비자 m 명에 확장한 형태로 확산모형 정의하였다. 이러한 Bass 모형은 F_t 값에 대한 비선형미분방정식의 해

를 구해서 연속적 시점에 대해 적용될 수도 있으나, 혁신수용자 수를 관측하는 이산적 시점에 맞춰 $S_t = (p+qY_{t-1}/m)(m-Y_{t-1})$ 로 정의되어 많이 적용되어 왔다. 이 때 S_t 와 F_t 는 각각 해당 혁신신제품에 대한 t 시점의 최초수용자의 수와 t 시점까지 누적된 최초수용자의 수를 의미하고, p , q 와 m 은 각각 혁신계수, 모방계수와 잠재시장규모를 나타내는 모수이다. Bass 모형에 대한 상세한 설명은 Mahajan and Muller(1979), Mahajan et al.(1990) 등을 참고할 수 있다.

Bass et al.(1994)은 Bass 모형에 여러 마케팅의 사결정변수의 영향을 포함시키기 위해 Bass 모형의 hazard 함수를 여러 마케팅 변수의 영향을 반영하기 위한 x_t 를 결합한(즉, $f_t/(1-F_t) = (p+qF_t)x_t$) GBM(generalized Bass model)을 제시하였다. Bass et al.(1994)의 GBM과 유사한 형태로 Joo and Jun(1996)은 시장환경에 따른 초기구매자수의 변동이 급격한 확산과정을 설명하기 위해 Bass 모형을 따르는 정상적인 확산과정과 시장환경에 따른 일시적 변동을 결합한 모형을 제시하였다. Joo and Jun(1996)은 Bass 모형의 이산식을 기준으로 t 시점의 최초수용자의 수(S_t)를 $(1+\delta_t)(p+qY_{t-1}/m)(m-Y_{t-1})$ 로 제시하고 있는데, δ_t 는 여러 시장환경요인들의 선형결합으로 표현되었다.

앞의 2.2절에서 고찰한 것과 같이 영화의 확산과정을 다룬 연구들 중에서도 Bass 모형에 여러 영향요인들을 결합한 연구들이 수행되어 왔다. 영화패널자료를 이용한 연구에서 Dellarocas et al.(2007)은 i 영화의 t 시점의 매출(Y_{it})을 Bass 모형의 이산식에 대응되는 형태로 제시하고, 혁신계수, 모방계수, 잠재시장크기 등의 모수는 여러 영향요인들의 선형회귀식으로 정의한 단계별 모형을 적용하였다. 이유석 외(2017), 차경천 외(2014), Lee et al.(2022) 등도 영화패널자료를 이용하여 영화관람객수의 확산을 GBM모형으로 제

시하고, 혁신계수, 모방계수, 잠재시장크기 등의 모수는 여러 영향요인들의 선형회귀식으로 정의한 단계별 모형을 적용하였다.

본 연구에서는 앞의 2.3절에서 제시된 연구문제들을 영화패널자료를 이용하여 분석하기 위해 영화확산에 대한 여러 영향요인들을 Bass 모형에 내재한 식 (1)의 Mixed-effects Bass 모형을 적용하고자 한다.

$$S_{i,t} = (p_{i,t} + q_{i,t} \frac{Y_{i,t-1}}{m_i})(m_i - Y_{i,t-1}) + \epsilon_{i,t} \quad (1a)$$

$$p_{i,t} = p_0 + d_1 D_{1,i,t} + p_1 Z_{1,i,t} + u_i^p \quad (1b)$$

$$q_{i,t} = q_0 + q_2 Z_{2,i,t} + q_3 Z_{3,i,t} + u_i^q \quad (1c)$$

식 (1a)는 i 영화의 t 시점의 관람객수를 설명하기 위한 Bass 모형의 이산식에 오차항($\epsilon_{i,t}$)을 더한 비선형식이고, 식 (1b)와 식 (1c) 각각 식 (1a)의 모수들이 각각의 영향요인들에 따라 결정되는 것을 나타내는 회귀식이다. 식 (1a)에서 $Y_{i,t}$ 는 i 영화의 t 시점까지의 누적관람객수를, $p_{i,t}$ 와 $q_{i,t}$ 는 각각 i 영화의 t 시점의 혁신계수와 모방계수를, m_i 는 i 영화의 잠재적 총관람객수를 나타낸다.

식 (1a)의 혁신계수($p_{i,t}$)에 대한 회귀식인 식 (1b)에서 D_1 은 영화별로 첫주를 나타내는 더미변수이고 $Z_{1,i,t}$ 는 i 영화의 t 시점의 스크린수(평균과의 차이)를 나타낸다. 식 (1b)에서 첫주를 나타내는 더미변수(D_1)를 포함한 것은 영화의 확산과정이 개봉전 홍보의 영향으로 발생된 큰 규모의 대기수요의 영향으로 첫주에 관람객 수가 매우 큰 것을 반영하기 위한 것이다. 영화확산과정을 설명하고 있는 기존 연구들 중에서도 김태구와 홍정식(2015)은 적분상수를 추가한 Bass 모형(BMIC: Bass model with integration constant)을 이용하여 확산초기에 영화 관람객 수가 큰 것을 반영하였다. 그리고 Bass 모형의 혁신계수는 기업이 소비자에게 전달하는 정보에 따른 소비자의 혁신수용

을 반영한 것이기에 영화의 확산과정에서도 기업의 의사결정이 반영된 스크린수(Z_1)를 혁신계수의 영향요인으로 고려하였다.

식 (1a)의 모방계수($q_{i,t}$)에 대한 회귀식인 식 (1c)에서 $Z_{2,i,t}$ 와 $Z_{3,i,t}$ 는 각각 i 영화의 t 시점의 리뷰수(평균과의 차이)와 평점(평균과의 차이)을 나타낸다. Bass 모형의 모방계수는 시장내에서 기존의 혁신수용자들이 발생시키는 구전(WOM: word of mouth)에 따른 소비자의 혁신수용을 반영한 것이기에 영화의 확산과정에서도 구전의 양(volume)을 반영한 리뷰수(Z_2)와 구전의 방향성(valence)을 반영한 평점(Z_3)을 모방계수의 영향요인으로 고려한 것이다.

식 (1b)와 식 (1c)에서 u_i^p 와 u_i^q 는 각각 $p_{i,t}$ 와 $q_{i,t}$ 에 대한 회귀식의 오차항을 나타낸다. 한편, 본 연구에서는 여러 영향요인들이 영화확산에 미치는 영향을 우선 혁신효과와 모방효과를 중심으로 분석하기 위해 식 (1a)에서 잠재적 총관람객수를 나타내는 m_i 는 각 영화별로 최종적인 누적관람객수로 알려진 상수로 취급하고, 이후 이를 추정대상인 모수로 확장한 분석을 추가하고자 한다.

식 (1)의 Mixed-effects Bass 모형은 이유석 외(2017), 차경천 외(2014), Dellarocas et al.(2007), Lee et al.(2022) 등에서 적용한 모형들과 같은 단계별모형(hierarchical model)의 구조를 지닌다. 식 (1a)는 영화 관람객수를 설명하기 위한 1단계모형이고, 식 (1b)와 식 (1c)는 식 (1a)에서 영화패널별로 지니는 이질성(heterogeneity)에 따른 패널효과를 1단계모형의 모수 $p_{i,t}$ 와 $q_{i,t}$ 를 통해 영화패널별로 반영하기 위한 2단계모형이다. 식 (1b)와 식 (1c)에서 상수항인 p_0 와 q_0 는 각각 모든 영화패널들이 공유하고 있는 혁신계수와 모방계수의 크기를 나타내고, 영화패널별로 다른 영향변수들인 $\{Z_1 \sim Z_3\}$ 와 영화패널별 오차항 $\{u^p, u^q\}$ 는 패널효과를 반영하게 된다.

식 (1)과 유사한 단계별 모형에 대해 이유석 외

(2017), 차경천 외(2014), Lee et al.(2022) 등은 각각의 영화들에 대한 확산과정을 추정하기 위한 1 단계모형과 혁신계수와 모방계수 등 확산과정의 모수의 영향요인을 회귀식으로 탐색하기 위한 2 단계모형을 명시적으로 구분한 방법을 적용하였다. 반면에 Dellarocas et al.(2007)은 단계별모형을 통합하여 추정하기 위해 MCMC(Markov Chain Monte Carlo)법을 적용하였다. MCMC법은 통계모형에서 추정해야 할 모수가 많아서 많은 모수를 한꺼번에 고려한 결합확률밀도함수(joint probability density function)로 정의되는 우도함수(likelihood function)를 적용하기 어려운 경우, 많은 모수들을 단계적으로 구분하여 정의한 조건부확률밀도함수(conditional probability density function)를 이용하여 추정대상 모수들에 대한 베이시안(Bayesian) 사전(prior)-사후(posterior) 확률구조를 적용한 시뮬레이션에 의한 추정방법이다.

한편, 단계별모형을 통합하여 추정하기 위해 MCMC법이나 사후확률이 명확히 정의 가능한 경우 적용될 수 있는 Gibbs 표본법 등의 Bayesian법 외에 ML(maximum likelihood)법이나 REML(restricted ML)법이 적용될 수도 있다. 단계별선형모형(HLM: hierarchical linear model)의 추정을 위한 ML법, REML법, Bayesian법 등에 대한 설명은 Raudenbush and Bryk(2002)을 참고할 수 있으며, 단계별비선형모형(hierarchical nonlinear model)의 추정을 위한 ML법과 Bayesian법 등에 대한 설명은 Davidian and Giltinan(2003)을 참고할 수 있다.

IV. 분석결과 및 해석

1. 분석결과

본 연구에서는 2.3절에서 제시한 연구문제를 탐

색하기 위하여 3.1절에 소개한 2018년부터 2021년까지 한국에서 개봉된 연도별 흥행실적 상위 30편씩 총 118편의 영화에 대한 관람객수와 스크린수(영화진흥원 통합전산망(<https://www.kobis.or.kr>) 제공), 온라인구전에 대한 리뷰수와 평점(다음영화사이트(<https://movie.daum.net>)에서 스크래핑) 등의 변수로 구성된 주별패널자료를 3.2절에 설명한 Bass 모형과 영향요인들을 결합하여 제시한 식 (1)의 단계별비선형모형(hierarchical nonlinear model)인 Mixed-effects Bass 모형을 이용하여 분석하였다.

본 연구에서 사용된 주별패널자료는 118편의 영화에 대해 영화별로 평균 9.907주(최소 3주, 최대 50주)로 모두 1,169개의 관측치로 구성된 불균형패널(unbalanced panel)이다. 본 연구에서 사용된 관람객수, 스크린수, 리뷰수와 평점 등에 대한 주요 기술통계량은 <표 1>에 제시하였다.

본 연구에서는 식 (1)의 Mixed-effects Bass 모형을 Stata의 menl(nonlinear mixed-effects regression) 명령어를 이용하여 추정하였다. Stata의 menl에서는 식 (1a)의 영화확산과정과 식 (1b)와 식 (1c)의 혁신계수와 모방계수에 대한 회귀식을 통합한 추정을 위해 ML(maximum likelihood)법이나 REML(restricted ML)법을 적용함으로써, 단계별 모형의 통합 추정을 위해 Dellarocas et al.(2007)에서 적용된 MCMC법과 구분될 수 있다. 이에 따른 식 (1)의 추정 결과는 <표 2>와 같다.

<표 2>에서 식 (1)의 Mixed-effects Bass 모형에 대한 추정 결과는 Model-3으로 표시된 열에 나타난 것이고, Model-1과 Model-2로 표시된 열은 식 (1)에 대한 모형적합성을 검토하기 위해 고려된 모형들에 대한 추정 결과이다. Model-1은 기본적인 Bass 모형으로 식 (1)에서 혁신계수와 모방계수에 대한 영향요인들과 오차항을 모두 제외하고 모든 영화들에 대한 혁신계수와 모방계수를 같은 값으로 설정함으로써 식 (1b)와 식 (1c)가 각각

<표 1> 주요 기술통계량

변수	총 관측치	평균	표준편차	최소값	최대값
관람객수	1,169	349,223	722,190	1	6,308,048
스크린수	1,169	3,197.7	3,315.4	1	17,925
리뷰수	1,169	141.28	296.57	1	2,877
평점	1,169	7.08	1.65	1	10

<표 2> 영화확산에 대한 Mixed-effects Bass Model 추정결과

		Model-1 (Bass)	Model-2 (without error in p & q)	Model-3 (Mixed-effects Bass)
		계수추정치, $z(p$ -값)	계수추정치, $z(p$ -값)	계수추정치, $z(p$ -값)
혁신 계수	p_0	0.320, 57.59(<0.001)	0.029, 2.07(0.038)	0.384, 23.12(<0.001)
	개봉첫주		0.202, 15.72(<0.001)	0.0110, 1.82(0.069)
	스크린수		2.06e-5, 16.19(<0.001)	7.36e-6, 5.07(<0.001)
모방 계수	q_0	0.251, 12.44(<0.001)	0.476, 18.30(<0.001)	0.284, 12.95(<0.001)
	리뷰수		3.49e-4, 11.77(<0.001)	3.29e-4, 23.05(<0.001)
	평점		-0.0479, -3.44(0.001)	3.19e-2, 3.98(<0.001)
RE	$\sigma_{u^2}^2$			0.0180
	$\sigma_{v^2}^2$			0.0329
	σ_{ϵ}^2	9.90e10	5.86e10	2.86e9
모형 적합성	MAPE	6.586%	6.265%	1.110%
	AIC	32,920.71	32,315.38	29,500.54
	BIC	32,935.90	32,350.82	29,546.11
	LR v.s. Model-1		$\chi^2_{df=4}=613.33$ (<0.001)	$\chi^2_{df=6}=3,432.17$ (<0.001)
	LR v.s. Model-2			$\chi^2_{df=2}=2,818.84$ (<0.001)

$p_{i,t} = p_0$ 과 $q_{i,t} = q_0$ 로 단순화된 모형이다. Model-2는 Model-1에 오차항 없이 혁신계수와 모방계수에 대한 영향요인들만을 추가함으로써 식 (1b)와 식 (1c)를 각각 $p_{i,t} = p_0 + d_1 D_{1,i,t} + p_1 Z_{1,i,t}$ 과 $q_{i,t} = q_0 + q_2 Z_{2,i,t} + q_3 Z_{3,i,t}$ 로 설정한 모형으로 Bass et al.(1994)의 GBM(generalized Bass model)에 대응되는 모형이다.

<표 2>에서 모형적합성을 위한 기준으로 제시된 MAPE(mean absolute percentage error)는 Model-1과 Model-2는 각각 6.586%와 6.265%인데 비해 Model-3는 1.110%로 모형의 잔차(residual)

가 Model-3에서 크게 감소하였고, AIC(Akaike information criterion)와 BIC(Bayesian information criterion) 모두 Model-1과 Model-2에 비해 Model-3가 낮게 나타나고 있다. 또한, 우도비(LR: likelihood ratio)검정 결과에서는 Model-2는 Model-1에 비해, Model-3은 Model-1과 Model-2에 비해 통계적으로 유의하게 설명력이 높은 모형인 것으로 나타나고 있다. 이에 따라 <표 2>의 모형 추정 결과들 중 Model-3의 Mixed-effects Bass 모형에 대한 추정 결과를 해석하는 것이 타당하다고 할 수 있다.

2. 해석

<표 2>에서 식 (1)의 Mixed-effects Bass 모형 (Model-3)에 대한 추정 결과는 다음과 같이 해석될 수 있다.

먼저, 개봉첫주, 스크린수, 리뷰수, 평점 등의 영향요인이 반영되지 않은 일반적인 혁신계수와 모방계수는 각각 0.384(s.d.=0.017)와 0.284(s.d.=0.022)로 나타났다. 영화확산과정에 대한 단계별 모형을 적용한 Dellarocas et al.(2007)은 혁신계수와 모방계수에 대한 회귀식의 상수항을 각각 0.035(s.d.=0.057)와 0.273(s.d.=0.087)로 보고하고 있고, 여러 영화들에 대한 확산과정의 혁신계수와 모방계수의 평균값을 Lee et al.(2022)은 각각 0.045(s.d.=0.023)와 0.047(s.d.=0.034)로, 이윅석 외(2017)는 각각 0.051(s.d.=0.026)과 0.055(s.d.=0.028)로 보고하였다. 일반적인 신제품의 확산모형을 다룬 연구들을 대상으로 한 메타연구에서 Sultan et al.(1990)은 혁신계수와 모방계수의 평균값을 각각 0.04와 0.302로 보고되었다. 단계별 모형의 혁신계수와 모방계수에 대한 식에 포함된 영향요인들을 본 연구에서는 식 (1b)와 식 (1c)에서와 같이 각 영향요인들의 평균과의 차이로 반영하는 등 각각의 연구들에서 정의된 식이 다르기 때문에, 본 연구의 혁신계수와 모방계수에 대한 추정치를 선행연구들에서 보고된 추정결과들과 직접 비교하는 것은 어렵다. 다만, 본 연구의 추정결과는 선행연구들의 추정결과에 비해 모방계수에 비해 상대적으로 혁신계수가 크게 나타났다. 이는 영화확산과정에서 혁신효과를 일으키는 기업의 정보전달노력의 중요성이 강조되어야 하는 것으로 해석될 수 있는데, 함수적으로는 영화확산이 종모양을 따르는 일반적인 혁신확산과정과 달리 확산초기부터 정점이 발생하고 이후 체감하는 모양을 나타내는 것을 반영한 결과이기도 하다.

혁신계수식(식 (1b))의 영향요인들인 개봉첫주와 스크린수의 계수추정치에 대한 단측검정을 위한 p -값은 각각 0.035와 0.001 미만값으로 개봉첫주와 스크린수 모두 유의하게 혁신계수에 정(+의 영향을 미친다. 개봉첫주의 혁신계수가 일반적인 혁신계수보다 더 큰 것은 위의 3.2절에서 모형에 대해 설명한 것과 같이 개봉전 홍보의 영향으로 발생된 큰 규모의 대기수요가 첫주 관람객 수로 전환된 것을 반영한 것으로 해석될 수 있다. 스크린수가 많을수록 혁신계수가 커지는 것은 스크린수의 많고 적음이 기업이 영화를 시장에 알리기 위한 노력 정도가 반영한 것으로 해석될 수 있다. 스크린수의 혁신계수에 대한 영향에 따라 영화개봉 초기에는 평균적인 수준보다 많은 스크린수에 의해 혁신효과에 의한 관람객수가 크게 나타나고, 영화가 개봉이후 시간이 지남에 따라 스크린수가 감소하며 해당 영화에 대한 혁신효과에 의한 관람객수는 급격히 감소하는 결과를 보이게 된다.

모방계수식(식 (1c))의 영향요인들인 리뷰수와 평점의 계수추정치에 대한 단측검정을 위한 p -값은 모두 0.001 미만값으로 리뷰수와 평점 모두 유의하게 모방계수에 정(+의 영향을 미친다. 특히, 평점의 계수추정치(0.0319)는 식 (1b)와 식 (1c)에서 오차항을 제외하고 추정된 <표 2>의 Model-2에 대한 추정치(-0.0479)가 유의한 음(-)의 영향이 있는 것과 대비되고 있다. 이는 <표 2>의 Model-2는 영화패널들이 지니는 패널효과들을 리뷰수와 평점의 차이만으로 한정하여 반영한 것이지만, 본 연구에서 적용한 Mixed-effects Bass 모형은 영화패널들이 지니는 패널효과들을 리뷰수와 평점뿐 아니라 그 외 많은 요소들을 오차항을 통해 반영하고 있는 차이에 의한 것이다. 4.1절에서 설명한 것과 같이 <표 2>의 모형적합성 기준들에 따라 가장 적합하게 평가되는 Model-3의 Mixed-effects Bass 모형을 적용할 경우 평점이 모방계수에 유

의한 정(+)¹⁾의 영향을 미치는 것으로 판단할 수 있는 것을, Model-2와 같이 패널효과가 불충분하게 고려된 모형을 적용할 경우 유의한 음(-)²⁾의 영향이 있는 것으로 다르게 판단할 수 있는 것이다.

3. 추가분석

지금까지 본 연구에서 <표 2>의 Model-3에 제시된 결과와 같이 추정된 식 (1)의 Mixed-effects Bass 모델을 중심으로 스크린수, 리뷰수, 평점 등이 영화확산과정에 미치는 영향을 분석한 결과는 영화패널별로 확산모형의 잠재적 총관람객수(식 (1)의 m_i)를 각 영화별 최종 누적관람객수로 알려진 상수로 두고 추정한 결과이다. 그러나 이는 영화확산과정에서 스크린수, 리뷰수, 평점 등의 영향요인들이 혁신계수와 모방계수를 통해 나타내는 혁신효과와 모방효과 외에 총관람객수에 미치는 효과를 제한하고 있는 한계를 가진다.

이에 본 연구에서는 추가분석을 통해 식 (1)의 잠재적 총관람객수(식 (1)의 m_i)를 상수가 아닌 추정대상 모수로 정하여 혁신계수식(식 (1b)) 및 모방계수식(식 (1c))과 유사하게 잠재적 총관람객수에 대한 회귀식을 추가한 모형에 대한 추정을 시도해 보았다. 이를 위해 식 (1)에 영화패널별로 잠재적 총관람객수를 첫주관람객수로 설명하는 식($m_i = m_1 \text{Fisrt Week Adnc}_i + u_i^m$)을 추가한 모형을 추정하였으며, 추정결과는 <표 3>과 같다.

<표 3>은 비교를 위해 식 (1)에 잠재적 총관람객수식을 추가한 모형에 대한 추정결과(Model-4)를 <표 2>에 나타난 식 (1)에 대한 추정결과(Model-3)와 함께 제시하였다. <표 3>에 의하면 영화별 잠재적 총관람객수는 첫주관람객수의 3.998배에 오차항이 추가되어 결정되는 것으로 나타났다는데, 실제 본 연구에서 분석한 118편의 영화들의 최종 누적관람객수는 평균적으로 첫주관람객수의 2.852배로 나타났다. 확산모형의 추정결과

<표 3> 영화확산에 대한 Mixed-effects Bass Model의 추가분석

		Model-3	Model-4
		계수추정치, z(p-값)	계수추정치, z(p-값)
혁신계수	p_0	0.384, 23.12(<0.001)	0.197, 25.87(<0.001)
	개봉첫주	0.0110, 1.82(0.069)	-0.0303, -7.72(<0.001)
	스크린수	7.36e-6, 5.07(<0.001)	1.07e-5, 11.59(<0.001)
모방계수	q_0	0.284, 12.95(<0.001)	-0.309, -13.92(<0.001)
	리뷰수	3.29e-4, 23.05(<0.001)	2.86e-4, 21.23(<0.001)
	평점	3.19e-2, 3.98(<0.001)	1.35e-3, 0.93(0.352)
잠재적 총관람객수	첫주관람객수		3.998, 22.72(<0.001)
RE	σ_{it}^2	0.0180	0.00256
	σ_{it}^2	0.0329	0.02602
	$\sigma_{u^m}^2$		8.75e12
	σ_{ϵ}^2	2.86e9	4.86e9
모형적합성	MAPE	1.110%	1.657%
	AIC	29,500.54	30,632.29
	BIC	29,546.11	30,687.99

잠재시장규모가 실제보다 과대추정이 되면 혁신계수와 모방계수는 이를 조정하기 위해 과소추정되는데, Bass 모형의 추정시 모수추정치들이 서로 민감하게 영향을 받는 것은 Tigert and Farivar(1981)에서 밝힌 바 있다. <표 3>에서와 같이 영화별 잠재적 총관람객수가 실제보다 과대추정됨에 따라 Model-4의 일반적인 혁신계수(p_0)와 모방계수(q_0), 개봉첫주 등에 대한 추정치는 Model-3에서의 추정치에 비해 작게 추정되었다. 그리고, Model-4에서 스크린수가 혁신계수에 미치는 영향과 리뷰수가 모방계수에 미치는 영향은 Model-3에서와 같이 유의한 정(+)의 영향으로 나타났지만, Model-4에서 평점이 모방계수에 미치는 영향은 Model-3에서와 달리 유의하지 않게 나타났다.

다만, <표 3>에서 AIC와 BIC를 기준으로 Model-4에 비해 Model-3의 모형적합성이 더 우수하게 나타나고, Model-4는 영화별 잠재적 총관람객수를 각 영화별 최종 누적관람객수에 비해 다소 과대추정하고 있다. 이에 본 연구에서는 스크린수, 리뷰수, 평점 등 여러 영향요인들이 영화확산에 미치는 영향에 대한 분석결과를 해석함에 있어, 위의 4.2절에서와 같이 영화별 잠재적 총관람객수를 최종 누적관람객수로 정한 Model-3의 추정결과를 이용하였다.

V. 결론

1. 결론 및 의의

본 연구에서는 영화확산과정에서 온라인구전의 영향력을 분석하기 위한 확산모형(diffusion model)으로 Mixed-effects Bass 모형을 제안하고, 제안된 모형을 한국영화시장에서 2018년부터 2021년까지 4년간 개봉된 연도별 박스오피스 상위 30편에 해당하는 총 118편의 영화를 대상으로 한 주별패널

자료에 적용하여 영화확산과정에서 온라인구전 양(volume)과 방향성(valence) 등의 영향력을 분석하였다.

본 연구에서 적용한 Mixed-effects Bass 모형은 확산과정을 설명하기 위해 가장 많이 사용되고 있는 Bass 모형(Bass, 1969)을 영화확산과정에 적용하되, Bass 모형의 모수인 혁신계수와 모방계수가 개봉첫주(더미변수), 스크린수, 리뷰수, 평점 등의 영향요인들(고정효과(FE: fixed effects))과 오차항(변동효과(RE: random effects))이 결합된 회귀식으로 설명되어지는 계층적비선형모형(hierarchical nonlinear model)의 형태를 지닌다. 분석결과 영화 확산과정에서 혁신계수는 개봉첫주와 스크린수에 유의한 정(+)의 영향을 받고, 모방계수는 리뷰수와 평점에 유의한 정(+)의 영향을 받는 것으로 나타났다. 이에 따라 영화확산과정은 개봉전 기대를 반영한 대기수요가 첫주에 혁신효과를 통해 나타나고(개봉첫주가 혁신계수에 미치는 정(+)의 영향), 기업의 노력 정도가 반영된 스크린수(스크린수가 혁신계수에 미치는 정(+)의 영향)와, 시장내에서 관람객들이 일으키는 구전의 양과 방향성을 반영한 리뷰수와 평점(리뷰수와 평점이 모방계수에 미치는 정(+)의 영향)에 따라 관람객수가 변화하는 것으로 이해될 수 있다.

본 연구의 의의는 다음과 같이 요약될 수 있다.

먼저, 2차자료를 이용하여 온라인구전의 영향력을 중심으로 영화성과(매출액 또는 관람객수)를 분석한 많은 선행연구들이 영화성과의 높은 왜도(skewness)를 고려하여 로그변환(log transformation)한 영화성과를 대상으로 선형모형을 이용하여 분석하고 있는 것에 대비하여, 모형의 추정과정에서 상대적으로 중요한 관측치라 할 수 있는 일부 극단적으로 높은 성과의 실질적인 가중치를 높일 수 있도록 로그변환을 하지 않은 성과변수 자체를 비선형모형으로 분류될 수 있는 Bass 모형을

기반으로 분석한 결과를 제시하고 있다. 또한, 본 연구는 기존 연구들 중 Bass 모델을 기반으로 분석하며 단계별모형을 MCMC(Markov Chain Monte Carlo) 방법으로 추정한 Dellarocas et al.(2007)과 단계별모형을 확산과정에 대한 추정과 확산모형 모수에 대한 회귀식을 통합하여 추정하는 대신 순차적으로 추정한 차경천 외(2014) 등과는 달리 비선형혼합모형(Mixed-effects nonlinear model)을 우도함수(likelihood function)를 기준으로 통합 추정한 결과라는 점에서 차별화되고 있다.

다음으로 본 연구에서는 Mixed-effects Bass 모형을 통해 분석한 결과 온라인구전의 양(volume)과 방향성(valence)이 모두 영화확산에 긍정적인 영향을 미치는 것으로 나타났다. 영화를 대상으로 온라인구전의 차원별 영향력을 분석한 선행연구들에서 분석에 사용된 모형이나 자료의 특성에 따라 온라인구전의 양과 방향성의 구전성과에 대한 영향력이 일관되지 못한 상황에서, 본 연구는 영화패널자료의 패널특성과 확산과정을 따른 추세적인 특성을 반영한 후 온라인구전의 양(리뷰수)과 방향성(평점)이 모두 영화확산에 긍정적인 영향을 미치는 것을 확인하고 있다는 점에서 의의를 찾을 수 있다.

그리고, 본 연구의 분석결과 개봉첫주의 큰 규모의 관람객수는 개봉전 형성된 대기수요와 스크린수의 영향을 받으며, 개봉첫주의 관람객수가 해당 영화의 총 누적관람객수에 유의한 영향을 미치는 것으로 분석되었다. 이는 영화성과를 높이기 위해 개봉전 프로모션과 개봉시 스크린수 확보 등을 통한 개봉첫주의 흥행실적 관리의 중요성을 시사하고 있다.

2. 한계점 및 향후 연구방향

본 연구의 한계점과 향후 연구방향은 다음과 같

이 요약될 수 있다.

먼저, 본 연구는 영화패널자료를 분석함에 있어 주별자료를 사용하였다. 이는 일별자료가 갖는 7일 주기의 계절성을 통제하며 일별자료를 사용하기가 매우 까다롭기 때문이다. 일별자료에 존재하는 7일 주기의 계절성을 통제하기 위해 이유석 외(2017), 차경천 외(2014), Lee et al.(2022) 등은 금요일, 토요일, 일요일 및 공휴일에 대한 더미변수를 사용하였으나, Dellarocas et al.(2007)은 본 연구에서와 같이 주별자료를 사용하였다. 다만, 영화확산이 초기 1~2주의 영향요인들에 의해 매우 민감하게 진행되는 과정에서 일별자료를 통해 좀 더 촘촘한 관찰주기를 분석하기 위해 향후 연구에서는 본 연구에서 적용한 Mixed-effects Bass 모형에 7일 주기의 계절성을 결합하기 위한 방법을 개발하여 일별자료를 대상으로 분석할 필요가 있다.

다음으로 본 연구에서는 분석에 적용한 Mixed-effects Bass 모형은 추가분석을 통해 잠재적 총관람객수를 추정한 결과를 제시하고는 있으나, 해당 결과에 대한 해석이 현실적인 결과인 최종 누적관람객수에 비해 잠재적 총관람객수를 크게 과대추정하고 있는 한계를 지니고 있다. 잠재적 총관람객수가 과대추정된 결과는 Tigert and Farivar (1981)에서 밝힌 것과 같이 Bass 모형의 추정시 모형의 모수에 대한 추정치들이 서로 민감하게 영향을 받고 있기 때문으로 해석될 수 있다. 이에 본 연구의 주된 분석은 Mixed-effects Bass 모형에서 잠재적 총관람객수는 영화별로 최종 누적관람객수로 알려진 상수로 두고 혁신계수와 모방계수에 대한 회귀식만을 적용한 모형을 적용하여 이루어졌다. 다만, 영화확산에 대한 여러 영향요인들의 영향력을 보다 종합적으로 이해하기 위해서는 이들 영향요인들이 혁신계수와 모방계수뿐 아니라 잠재적 총관람객수에 미치는 영향을 포함한 분석이 요구되기에 향후 연구에서는 본 연구에서

적용한 Mixed-effects Bass 모형을 현실과의 적합성이 높은 잠재적 총관람객수를 함께 추정할 수 있도록 개선할 필요가 있다.

마지막으로 본 연구에서 적용한 Mixed-effects Bass 모형에서는 영화확산에 영향을 줄 수 있는 많은 요인들 중 개봉첫주(더미변수), 스크린수, 리뷰수 및 평점 등의 4가지 요인들만을 제한적으로 포함하고 있다. 이는 부분적으로 본 연구에서 사용한 주별자료가 영화별로 관측치가 충분하지 않아서 모형에 투입되는 영향요인의 수가 다소 제한적이기 때문이다. 향후 연구에서는 앞서 인식한 일별자료를 대상으로 한 분석모형 등을 통해 영화별로 충분한 관측치를 포함하여 장르, 제작국가, 감독, 배우 등 다양한 영화특성요인들 및 상영시점의 경쟁영화특성요인들의 영향력을 함께 분석할 필요가 있다.

논문접수일: 2023. 09. 22.

1차 수정본 접수일: 2023. 10. 12.

게재확정일: 2023. 10. 17.

참고문헌

- 김태구, 홍정식 (2015). 한국과 미국에 있어 영화 수익관련 통계량과 확산 현상의 비교 분석. *경영과학*, 32(1), 133-145.
- 류강석, 유시진 (2015). 구전연구에 관한 문헌적 고찰: 2004-2014 국내 마케팅 학술지를 중심으로. *마케팅연구*, 30(1), 145-176.
- 배정호, 심범준, 김병도 (2010). 온라인 구전과 영화 매출 간 상호영향에 관한 연구: 한국 영화 산업을 중심으로. *아시아마케팅 저널*, 12(2), 1-25.
- 이유석, 차경천, 김상훈 (2017). 배급사의 의사결정이 영화 확산에 미치는 영향: 한국시장의 사례를 중심으로. *마케팅연구*, 32(3), 25-44.
- 이중원, 박철 (2019). 온라인 구전의 방향성과 분산이 영화매출에 미치는 영향: 경쟁영화 온라인 구전 특성의 조절효과를 중심으로. *경영학연구*, 48(2), 341-360.
- 장광필 (2008). 영화 개봉 전 프로모션과 개봉 후 구전효과에 관한 연구. *마케팅관리연구*, 13(4), 23-34.
- 주영진, 최유리 (2021). 패널오차수정모형을 이용한 온라인 구전이 영화 관객수에 미치는 영향 분석. *경영과학*, 38(4), 49-66.
- 주영진, 최유리 (2023). 영화확산에 대한 온라인 구전의 차원 및 영향력 분석. *경영과학*, 40(2), 125-145.
- 차경천, 천정빈, 윤성욱 (2014). 구전자산: 측정법과 영화산업에의 응용. *마케팅연구*, 29(1), 179-195.
- Bass, F. M. (1969). A new product growth for model consumer durables. *Management Science*, 15(5), 215-227.
- Bass, F. M., Krishnan, T. V., & Jain, D. C. (1994). Why the bass model fits without decision variables. *Marketing Science*, 13(3), 203-223.
- Chintagunta, P. K., Gopinath, S., & Venkataraman, S. (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.

- Davidian, M. & Giltinan, D. M. (2003). Nonlinear models for repeated measurement data: An overview and update. *Journal of Agricultural, Biological, and Environmental Statistics*, 8(4), 387-419.
- Dellarocas, C., Zhang, X., & Awad, N. F. (2007). Exploring the value of online product reviews in forecasting sales: The case of motion pictures. *Journal of Interactive Marketing*, 21(4), 23-45.
- Duan, W., Gu, B., & Whinston, A. B. (2008a). Do online reviews matter?-An empirical investigation of panel data. *Decision Support Systems*, 45(4), 1,007-1,016.
- Duan, W., Gu, B., & Whinston, A. B. (2008b). The dynamics of online word-of-mouth and product sales—An empirical investigation of the movie industry. *Journal of Retailing*, 84(2), 233-242.
- Horsky, D. & Simon, L. S. (1983). Advertising and the diffusion of new products. *Marketing Science*, 2(1), 1-17.
- Joo, Y. J. & Jun, D. B. (1996). Growth-cycle decomposition diffusion model. *Marketing Letters*, 7(3), 207-214.
- Kim, K., Yoon, S., & Choi, Y. K. (2019). The effects of eWOM volume and valence on product sales – an empirical examination of the movie industry. *International Journal of Advertising*, 38(3), 471-488.
- Lee, Y., Kim, S., & Cha, K. C. (2022). The diffusion pattern of new products: evidence from the Korean movie industry. *Asian Business & Management*, <https://doi.org/10.1057/s41291-022-00196-0>.
- Liu, Y. (2006). Word of mouth for movies: Its dynamics and impact on box office revenue. *Journal of Marketing*, 70(3), 74-89.
- Mahajan, V. & Muller, E. (1979). Innovation diffusion and new product growth models in marketing. *Journal of Marketing*, 43(4), 55-68.
- Mahajan, V., Muller, E., & Bass, F. M. (1990). New product diffusion models in marketing: A review and directions for research. *Journal of Marketing*, 54(1), 1-26.
- McKenzie, J. (2022). The economics of movies (revisited): A survey of recent literature. *Journal of Economic Surveys*, DOI: 10.1111/joes.12498.
- Raudenbush, S. W. & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Sage Publications.
- Sultan, F., Farley, J. U., & Lehmann, D. R. (1990). A meta-analysis of applications of diffusion models. *Journal of Marketing Research*, 27(1), 70-77.
- Tigert, D. & Farivar, B. (1981). The Bass new product growth model: A sensitivity analysis for a high technology product. *Journal of Marketing*, 45(4), 81-90.
- You, Y., Vadakkepatt, G. G., & Joshi, A. M. (2015). A meta-analysis of electronic word-of-mouth elasticity. *Journal of Marketing*, 79(2), 19-39.

Modeling Movie Diffusion Using Mixed-effects Bass Model: Focusing on the Effect of eWOM

Youngjin Joo*

ABSTRACT

Purpose and Background: In this study, we analyzed the effects of a few key influencing factors, including eWOM (or online word-of-mouth), on the movie diffusion. One of the main interests in many previous researches aiming to analyze the effects of eWOM in the movie industry is comparing the effects of the two key dimensions (volume and valence) of eWOM, but the analysis results vary depending on the data and models used for the analysis.

As the use of secondary data, combining the movie box office data and the web-scraping review data from online movie websites, has become more accessible, there has been an increasing number of studies attempting to explain the movie performance (the movie box office revenue or the number of audiences) using those secondary data. And the main focuses of those studies are the effects of the number of reviews (or the volume of the eWOM) and the average rate of reviews (or the valence of the eWOM) on the movie performance. However, unlike the primary data, where the researcher can control the measurement units of variables according to a predefined research design, such secondary data possesses various panel heterogeneity and time series characteristics that should be effectively handled by the research model.

And, in many previous studies analyzing movie performance using secondary data, it is a common practice to analyze log-transformed movie performance variable and to apply a linear model, because the movie performance variables are highly skewed due to the exceptionally high revenue or number of audience of a few blockbuster movies. In order to increase the explaining power of the research model, it may be necessary to analyze movie performance variable itself, without log-transformation, using a nonlinear model. Some previous study has applied the Bass model, which can be categorized as a nonlinear model, but relatively more studies have used linear regression models.

Research model, data, and results: In this study, we applied the Mixed-effects Bass model to analyze the effects of factors such as the number of screens, the number of reviews and the average rate of reviews on the movie diffusion measured by the movie performance (the number of audience) itself without applying log-

* Professor, School of Business, Chungbuk National University

transformation. The Mixed-effects Bass model effectively captures the panel heterogeneity and time series characteristics of the movie panel data. It incorporates the primary long-term diffusion process, an inherent characteristic of the data, through the diffusion process inherent in the Bass model. The parameters of the Bass model, the innovation coefficient and the imitation coefficient, with regression models, are explained by both fixed effects (such as the number of screens, the number of reviews, and the average rate of reviews) and random effects (represented by error terms). These two levels (Bass model for movie diffusion and regression models to explain the Bass model parameters) are creating a hierarchical non-linear model.

In this study, we are using weekly panel data for a total of 118 movies released in the Korean film market from 2018 to 2021. These movies were selected based on the top 30 box office rankings each year. In the analysis using the Mixed-effects Bass model, in addition to the factors of the number of screens, the number of reviews, and the average rate of reviews, an opening week dummy variable reflecting pre-release expectations was added as an influencing factor for the innovation coefficient and imitation coefficient.

The analysis results indicated that the innovation coefficient in the Bass model of the movie diffusion is significantly positively influenced by the opening week and the number of screens, and the imitation coefficient is significantly positively influenced by the number of reviews and the average rate of reviews.

Conclusions: The implications of this study can be summarized as follows:

Firstly, this study represents the first application of the Mixed-effects Bass model to analyze the effects of factors such as eWOM on movie performance (the number of audience) using weekly panel data. This approach integrates the estimation of the Bass model with the estimation of regression equations. Where the Bass model represents the movie diffusion, and regression equations capture the effects of both fixed effects (such as the number of screens, the number of reviews, and the average rate of reviews) and random effects (represented by error terms) on the innovation coefficient and the imitation coefficient of the Bass model.

Additionally, the results of this study confirmed that both the volume of eWOM (the number of reviews) and the valence of eWOM (the average rate of reviews) have statistically significant effects on movie performance (the number of audience).

Keywords: eWOM, Movie Diffusion, Mixed-effects Bass Model, Hierarchical Nonlinear Model, Movie Panel Data