

A Study on the Content of Online Reviews and the Sales Performance: Based on the Tweets for Book Reviews

온라인 리뷰의 내용과 제품의 판매 성과에 관한 연구: 책 리뷰에 관한 트윗을 바탕으로

Jungmin Son¹

손정민¹

¹ Associate Professor, School of Business, Chungnam National University, South Korea,
sonjm@cnu.ac.kr

Abstract: Online review for products by users is an significant factor which could examine the sales performance of products. Prior researches focus on the volume of online reviews, valence, emotion, and social dimentions. Recently other theoretical dimensions are proposed for content analysis, such as the engagement and uniqueness seeking of reviewers and users. This study classified the content of the review into idea-sharing and information-sharing. Also, I analyzed the differential affect of thew two types on product performance. The idea-sharing type is to elicit elaborated information process to reviewers in the writing process. On the stage of the empirical analysis, to answer these research questions, I gather the sales data and users' reviews on Twitter of top 100 books in online bookstores. The data were collected on a daily level from January 2022 to March 2022. The data include product characteristics, brand performance of book authors, content of reviews, and Twitter user's characteristics. Based on the estimation results, idea-sharing reviews have more impact on the product performance than information-sharing reviews. In addition, the carryover effect of the review effect is significant, indicating that reviews have a significant impact on sales performance over a long-term. This research explain the implications for online retailers, product manufacturers, and reviewers. It suggests that online retailers need to promote more contributions and engagements to users who can more contribute idea-sharing reviews.

Keywords: Online Review, Idea Sharing, Online Book Store, Twitter, MSM(method of simulated moment)

요약: 온라인 리뷰는 제품의 판매 성과를 설명할 수 있는 중요한 요인이다. 이전의 전통적인 연구에서는 온라인 리뷰의 양, 긍정과 부정, 감정적 요소, 사회적 요소 등에 집중하여 연구가 수행되어 왔다. 최근에는 리뷰 작성자 및 이용자의 몰입, 차별성 등과 같은 리뷰의 내용을 분석할 수 있는 다양한 이론적 차원이 제시되고 있다. 이 연구는 리뷰의 내용을 아이디어 공유형과 정보 공유형으로 구분하고 판매 성과에 미치는 차별적인 영향을 분석하였다. 아이디어 공유형은 정보 공유형에 비해, 리뷰 작성 과정에 더 정교한 노력과 시간이 소요될 가능성이 높다. 이러한 연구 질문을 실증적으로 분석하기 위하여, 온라인 서점의 100개의 베스트셀러 도서의 판매 데이터 및 트위터 상의 리뷰 데이터를 2022년 1월부터 3월까지 일 단위로 수집하였다. 데이터에는 도서의 가격 정보와 같은 특성, 도서 저자의 브랜드 성과,

Received: January 25, 2023; 1st Review Result: March 12, 2023; 2nd Review Result: April 09, 2023
Accepted: April 30, 2023

리뷰의 내용, 리뷰 작성자의 팔로워 수 등과 같은 다양한 변수가 포함되었다. 이 연구의 분석 결과에 따르면, 아이디어 공유형 리뷰가 정보 공유형 리뷰에 비해 판매 성과에 더 큰 영향을 주었다. 또한 리뷰 효과의 이월 효과가 유의미하게 나타남으로써 리뷰가 장기간에 걸쳐 판매 성과에 유의미한 영향을 줄 수 있는 것으로 나타났다. 분석 결과를 바탕으로, 온라인 소매점과 제품의 제작자, 그리고 리뷰 작성자에게 실무적 시사점을 제공하고자 하였다. 특히 온라인 소매점이 아이디어 공유형 리뷰와 같이 몰입할 수 있는 이용자의 기여를 유도할 필요가 있음을 시사한다.

핵심어: 온라인 리뷰, 아이디어 공유, 온라인 서점, 트위터, MSM(method of simulated moment)

1. 서론

온라인 소매 시장의 규모가 커진데 비해 온라인 소비자들의 구매 행동에 영향을 주는 온라인 리뷰의 내용을 분석한 연구는 부족한 실정이라고 할 수 있다. 온라인 유통 산업이 발전하면서 일상에서 구매하는 제품중 온라인 채널을 이용하는 소비자가 많아졌다. 한국의 소매업 분야를 기준으로 온라인쇼핑 거래액이 200조원을 넘어섰으며[1], 이는 오프라인과 비슷한 규모로 성장한 것이다[2]. 시장의 성숙에 비해 제품 리뷰에 대한 질적인 이해는 이전 연구에서 비교적 부족한 편이다.

온라인 제품 리뷰에 관한 이전 연구는 리뷰의 양적인 요소에 보다 집중 해왔다[3]. 온라인 리뷰의 양, 참여자의 특성, 기업 가치에 미치는 영향 등에 관한 이전 연구에서는 리뷰의 중요성에 대한 이해를 높이는데 기여했다고 볼수 있을 것이다. 이용자들이 자발적으로 참여하여 생산된 리뷰 또는 구전(word-of-mouth)의 내용에 관한 연구분야가 최근 활발히 진행되고 있는데, 내용의 긍정 또는 부정의 비율, 그리고 감정 및 사회적 요소가 얼마나 포함되었는지에 따른 차이가 존재하는 것으로 보고 되었다[4][5]. 또한 리뷰를 작성하는 환경에 따라서 내용의 길이와 깊이가 달라지며[6], 이용자들이 새로움을 추구하는 현상으로 인하여 리뷰 작성에 노력을 얼마나 투입하였는지에 의해 사람들의 반응이 차별적으로 나타나는 것으로 알려져 있다[7][8].

이 연구의 목적은 제품 리뷰의 내용을 아이디어 공유형과 정보 공유형으로 구분하여 제품 성과에 미치는 차별적인 영향력을 장기적으로 설명하는 것이다. 정보 공유형의 제품 리뷰는 작성 과정에서 작성자의 정교한 언어처리 과정을 거쳤을 가능성이 있다[9]. 아이디어를 작성하는 과정은 정교한 언어를 처리하고, 노력과 시간을 상대적으로 많이 투입해야할 가능성이 있기 때문에[10][11], 새로운 내용을 추구하는 다른 이용자들에게 참여와 관심을 이끌어 낼수 있는 것으로 예상할 수 있다[8]. 또한 이 연구에서는 이용자가 작성한 제품에 대한 리뷰가 장기적으로 제품 성과에 미치는 영향에 대해 설명할 수 있는 모형을 제시하고 실제 온라인 플랫폼에서 측정된 데이터를 이용하여 분석하고자 한다. 이를 위하여 온라인 대형 서점의 판매 데이터 및 트위터 상에서의 도서에 대한 온라인 리뷰 및 이용자의 반응 데이터를 수집하여 분석하였다.

이 연구는 이전의 온라인 성과와 관련한 연구에 기반하고 있으며, 다음과 같은

차별적인 기여점을 보이고자 한다. 첫째, 리뷰의 질적인 차이에 의한 제품 성과 효과를 설명하는 것이다. 이 연구는 온라인 제품 판매 데이터에 미치는 리뷰 내용의 질적인 차이를 설명한 매우 드물고 중요한 연구가 될것이다[9]. 둘째, 이용자생산콘텐츠 연구

분야에서 장기적인 제품 성과의 변화를 설명할 수 있는 모형을 제시한 기여점이 있다. 이전 연구[12]에서는 브랜드 가치를 장기적으로 설명하기 위한 시도가 있었으며, 외부의 평가 요인에 의한 장기적인 브랜드 가치의 상승과 하락을 설명하였다. 이 연구는 해당 분야의 이전 연구에 기반하고 있으며 이용자생산컨텐츠와 같이 다른 이용자의 반응이 빠른 제품 카테고리도 연구 분야를 확장했다는 발전적 차별점이 있다. 셋째, 이 연구는 실제 온라인 서점과 외부 소셜네트워크인 트위터 데이터를 결합함으로써, 이용자생산컨텐츠와 온라인 상점의 제품 판매 성과의 관계를 실증적으로 분석한 기여점이 있다. 방대한 실증 데이터를 분석하여 리뷰의 질적인 차이를 설명하고자 하였다. 분석 데이터인 온라인 서점은 해당 국가에서 1위 업체를 대상으로 하였으며, 트위터 또한 온라인 소셜 미디어 상에서 매우 잘 알려진 플랫폼으로, 분석 결과의 일반화 가능성을 높였다.

2. 이론적 배경

2.1 온라인 제품 성과와 제품 리뷰의 관계

이전 연구에서 온라인 성과는 제품의 판매, 이익, 그리고 판매와 이익에 영향을 줄 수 있는 구매 의도 등 다양한 변수를 이용하여 측정해왔다[3][6][13]. 이와 같은 온라인 플랫폼에서의 종속 변수 측정에 관한 이전 연구에서는 시점별 개별 제품의 판매량을 분석할 수 있었다. 제품 판매량은 온라인의 상점의 성과에 관한 이전 연구[8]에서 주로 이용되었던 변수이다. 또한 온라인 리뷰와 온라인 제품 판매 성과의 관계는 이전 연구에서 온라인 서점, 온라인 유아용품 등과 같은 다양한 데이터 및 계량적 모형으로 설명 되어왔다[5][14][15]. 이전 연구 결과에 따르면 제품의 인기도는 제품의 판매 성과에 영향을 줄 수 있는 주요한 변수로 설명되어 왔다. 제품 인기도가 높을수록 이후의 다른 소비자의 선택 확률이 증가한다. 다만 온라인 채널은 오프라인에 비해 제품의 선택 대안이 많기 때문에 인기도가 낮은 제품을 소비자가 검색을 통하여 선택할 수 있는 확률이 높아지는 효과가 발생한다[14]. 따라서 인기도가 상대적으로 낮은 제품도 검색 엔진의 기술적 발전, 다른 이용자의 리뷰 생산 및 공유 활동으로 인하여 판매 성과를 얻을 수 있다. 즉 낮은 인기도의 제품의 선택이 온라인 소매점 매출의 상당 부분을 차지하는 롱테일(long-tail) 효과가 발생하는 특성이 있다. 온라인 리뷰 및 구전 등과 같은 이용자 및 전문가 등의 평가 활동을 온라인 채널의 제품 성과 분석에 포함해야 하는 중요한 이유라고 할 수 있다[3].

이 연구는 온라인 상점에서 시간에 따른 제품 성과를 설명하고자 한다. 이와 관련하여, 이전 연구에서는 온라인 플랫폼에서의 제품 성과는 이용자들 사이에 연결된 외부의 소셜 네트워크에 의해 중요한 영향을 받는다는 점을 설명해왔다[4][11]. 온라인 플랫폼은 이용자 사이에 연결이 쉽고 서로 아이디어와 다른 정보를 공유하는 데에 있어 시간 및 공간적 비용이 상대적으로 적다. 일반적으로 익명의 이용자 사이에는 정보의 전달 속도가 서로 얼굴을 아는 경우에 비해 더 빠르기 때문에[14], 외부 소셜네트워크에서의 정보 공유에 의한 제품 성과의 변화를 중요한 요인으로 살펴보아야 할 필요가 있을 것이다.

온라인 플랫폼에서는 소비자 사이의 연결이 용이한 특성으로 인하여 정보 전달의 속도가 오프라인에 비해 상대적으로 더 빠르다고 할 수 있다. 이러한 정보 전파 속도를 더욱 가속하는 요인은 정보의 감정적 특성이다[4][6]. 긍정 또는 부정적 감정중에서

감정의 강도가 강할수록 전파 속도가 더 빠르다고 알려져 있다. 따라서 구체적인 감정적 특성 및 강도가 포함되어 있는 텍스트 기반의 온라인 리뷰의 내용의 중요성이 높다[16]. 이 연구는 이러한 이전 연구를 기반으로 하여, 온라인 리뷰와 제품의 성과 관계를 장기적으로 추적 및 분석하고자 한다.

2.2 이용자생산콘텐츠

이용자생산콘텐츠(User Generated Content)는 이용자가 자발적으로 생산에 참여한 콘텐츠를 말한다. 온라인 플랫폼에서는 사용자들의 참여가 비교적 용이하며, 다른 사용자들의 평가를 수용하고 공유하는 행위가 널리 퍼져있다[9][12][14]. 이때 이용자의 리뷰의 효과는 리뷰 작성자의 팔로워수, 이전에 작성한 리뷰의 수, 인기도, 그리고 온라인 플랫폼 상에서 리뷰의 재 인용 빈도, 다른 사용자들의 반응, 그리고 리뷰를 작성한 시점에 따른 이월효과[12] 등에 의해 설명할 수 있다. 제품의 특성, 리뷰의 특성, 리뷰 작성자의 특성, 그리고 각 특성 요소의 효과 및 효과의 시간에 따른 이월 효과 등이 이용자생산콘텐츠의 성과를 설명하는 모형에 포함할 필요가 있는 것이다.

이용자의 참여는 단순한 정교 교환, 질문과 답변, 기업과의 피드백, 전문지식 및 콘텐츠의 공동 생산과 공유 등 다양한 형태와 수준에서 발생할 수 있다[12]. 사용자들의 참여는 기업이 소비자 관계 관리 차원에서 추구하는 만족을 넘어선 이용자의 몰입(engagement)을 유도한다는 점에서 매우 중요한 활동이라고 할 수 있다[17]. 이용자의 몰입은 기업의 제품 및 서비스와 함께하는 시간이 길어진 이후에 직접 참여하는 활동으로써, 이전 연구에서는 몰입의 측정을 위하여 이용자가 좋아요를 클릭한 횟수, 리뷰의 길이 및 작성 내용의 유형 등을 포함하기도 하였다[18][19].

이용자생산콘텐츠는 다른 사용자들의 리뷰와 같은 반응을 실시간으로 작성하여 공유할 수 있다는 특징이 있다. 따라서 제품의 성과는 이전의 제품에 대한 다른 사용자들의 반응에 의해 큰 영향을 받을 수 있다. 최종 판매 성과의 변화는 여러 시간대에서 제품에 대한 온라인 상에서의 리뷰에 의해 영향을 받을 가능성이 있다는 것이 알려져 있다[20]. 온라인 리뷰는 이용자의 자발적인 참여로 이루어지는 과정에서 발생하기 때문에, 리뷰의 효과 역시 시간에 따른 작성자의 활동과 특성에 영향을 받을 수 있다. 이전 연구에 따르면 이용자들이 자발적으로 참여하여 생산된 이용자생산콘텐츠의 품질은 언어적인 복잡성과 같은 리뷰 내용의 가독성, 정보의 길이와 내용을 조사함으로써 생산과정에 소모된 콘텐츠 생산자의 인지적 자원을 판단할 수 있다[17][18]. 또한 리뷰에는 텍스트 정보 이외에도 사진과 영상, 그리고 위치 등에 관한 정보가 포함되어 있으므로, 다양한 차원의 이용자 생산 정보에 대한 평가 체계를 이해할 필요가 있을 것이다.

2.3 온라인 제품 성과 측정 모형

이 연구의 목적인 온라인 리뷰 내용의 유형에 따른 제품 성과에 대해 설명하기 위해, 계량 경제학 및 경영학 분야에서의 이전의 관련 연구에서 이용해왔던 모형 구축 과정에 기반하고자 한다. 개별 제품의 판매 성과가 최종적인 결과 변수일때, 제품의 특성 및 리뷰의 변동성이 원인 변수로 설정되어 모형을 구축할 수 있다[20]. 특히 온라인 리뷰는 실시간으로 업데이트되며 사용자들의 반응 또한 장기간에 걸쳐서 누적적으로 발생한다는 특징이 있다[12]. 이러한 리뷰 및 온라인 판매 성과의 특성을 분석하기 위하여 변수들

사이의 정교한 관계를 장기간에 걸쳐 추정할 수 있는 MSM (method of simulated moment)과 같은 동시적 추정 방법이 활용되어 왔다[20]. 이외에도 브랜드 가치의 누적 모형, 이월효과 등과 같은 계량경제학 및 경영학 분야의 모형[21-23]을 도입하여 분석하였다.

이상에서 설명한 이론적 배경 및 관련 문헌을 아래와 같은 표로 나타내었다. 이 연구와 관련된 선행연구에 관한 도표로써 이 연구의 주요 변수를 중심으로 요약 및 정리하였다.

[표 1] 문헌 연구

[Table 1] Literature Review

| 논문 | 연구 주제 | 연구 범위 및 독립변수 | | | | | 종속변수 |
|---|----------------------------|--------------|-------|--------|-------|-------|-------------|
| | | 제품 특성 | 제품 리뷰 | 생산자 특성 | 이월 효과 | 동적 특성 | |
| J. Chevalier, D. Mayzlin (2006). | 도서 판매, 제품 가격, 구전 내용 사이의 관계 | ○ | ○ | | | | 도서 판매순위 |
| C. Kim, F. J. Costello, K. C. Lee (2020) | 게이미피케이션에 의한 새로움 추구 충족 효과 | ○ | | | | | 재구매 의도 |
| J.H. Chung and Y.C. Kim (2004) | 구전에 의한 영화 선택 효과 | ○ | ○ | ○ | | | 영화 선택 |
| A. Ghose, P. G. Ipeiros, B. Li. (2012) | 온라인 호텔 리뷰와 소비자의 선호도 관계 | ○ | ○ | | ○ | | 호텔 선호도 |
| U. Narang, M. S. Yadav, A. Rindfleisch (2022) | 아이디어 공유형 리뷰에 의한 이용자 참여 효과 | ○ | ○ | ○ | | | 이용자 몰입 |
| S. Melumad, J. J. Inman, M. T. Pham (2019) | 스마트폰과 PC의 UGC 내용의 차이 | ○ | ○ | | | | 리뷰의 내용 |
| L. Luo, X. Chen, J. Han, C. W. Park (2010) | 연예인의 영화 출연과 장기적 인기도 변화 모형 | ○ | | ○ | ○ | ○ | 연예인의 인기도 점수 |
| 본 연구 | 도서 리뷰의 내용과 판매 성과의 장기적 관계 | ○ | ○ | ○ | ○ | ○ | 도서 판매량 |

3. 연구 모형 및 데이터

앞서 이론적 배경에서 살펴본 바와 같이, 이 연구는 이용자의 리뷰를 두 유형의 정보 공유로 분류하였다. 아이디어 공유와 제품 정보 공유라는 두 유형으로 분류 하여 설명하고자 한다. 따라서 모형에는 소셜네트워크 상에서 공유되는 정보의 내용을 두 유형에 해당되는 정도를 주요한 독립 변수로 포함하였다. 이러한 공유 정보는 정보 생산자의 특성에 의해 영향을 받기 때문에, 정보 생산자의 팔로워, 좋아요, 재공유 등과

같은 정보 생산자의 특성을 고려하고자 한다. 이러한 특성의 효과는 외부 변인이 간섭하지 않을 경우 시간에 따라 점차 감소할 가능성이 있기 때문에, 이월효과를 고려하고자 한다. 이외에도 판매 제품의 특성 및 시간에 의한 환경 변화 등을 설명변수로 포함하고자 한다.

3.1 제품 성과 모형

이 연구에서는 개별 제품의 판매 성과를 시간에 따라 변화하는 정도에 대해 설명하고자 한다. 제품의 판매 성과는 제품의 이전 판매 성과 및 제품 자체의 품질 등과 같은 특성 요소에 의해 설명 할수 있는 부분이 있다[14][18].

이 연구가 제안하고자하는 제품 성과 모형은 다음의 식(1)과 같다. 해당 식에서 제품 i 의 t 시점에서의 판매량은 $Sales_{it}$ 로 나타낸다. 아래 모형에서는 시간에 따른 개별 제품의 판매량은 제품과 관련된 리뷰의 내용 및 리뷰의 작성자의 특성에 의해 설명 될수 있다고 가정한다. $Product_{ik}$ 는 제품 i 에 대한 k 시점에서의 특성에 대한 벡터로, 제품의 제작자 및 제품 자체의 특성에 대한 정보가 포함되어 있다. $Review_{ip}$ 는 제품 i 에 관한 리뷰 p 에 포함된 내용 중 아이디어 공유에 관한 내용의 비율 값이다. 이를 통해, 리뷰의 유형이 성과에 미치는 효과는 시간에 따라 변화하는 동태적 스케일을 통해 변화하는 과정을 설명 할 수 있다. 이에 해당하는 스케일 함수는 τ_{ikp} 로, 제품 및 리뷰별로 동적 스케일 값을 산출하여 장기간에 걸친 리뷰의 효과를 포착할 수 있다.

$$Sales_{it} = \sum_k^{t-1} \{ (\gamma_i \cdot Product_{ik} \cdot d_{1,it} + \tau_{ikp} \cdot Review_{ikp} \cdot d_{2,itp}) \cdot \prod_{m=1}^{t-1} \delta_{im} \} + Sales_{ij} \cdot \prod \delta_{ij} + \epsilon_{it} \quad (1)$$

위 식에 포함된 각 항과 계수에 대한 정의는 아래와 같다.

γ_i = 제품 i 의 효과 계수

$Sales_{ik}$ = 제품 i 의 시점 k 에서의 판매 성과

$Product_{ik}$ = 제품 i 의 시점 k 에서의 특성에 대한 벡터항

τ_{ikp} = 제품 i 에 대한 리뷰 p 의 시점 k 에서의 동적 스케일

$Review_{ip}$ = 제품 i 에 대한 리뷰 p 의 내용중 아이디어 공유의 비율

$d_{1,it}$ = 제품 i 의 시점 t 에서의 판매 여부에 대한 더미 변수

$d_{2,it}$ = 제품 i 에 대해 t 시점에 리뷰가 작성된 여부에 대한 더미 변수

δ_{im} = 리뷰 i 가 시점 t 에서 m 사이 동안의 이월효과 함수

ϵ_{it} = 정규분포를 따르는 오차항으로 구성된 판매의 기저선

장기간에 걸친 리뷰의 효과를 설명하기 위한 이전의 연구들에서는 성과에 영향을 주는 주요한 요소들이 장기간에 걸쳐 누적되는 효과가 있다면, 그 효과의 크기를 외생변수를 고려한 일정한 체계적 불확정성의 정도로 평가하고자 하였다[20]. 이 연구의 모형에서도 이러한 체계적 불확정성의 크기를 고려할수 있는 동적 특성 스케일 (i.e., τ_{ikp})을 포함하였다. 동적 특성 스케일은 비율값을 가질 수 있도록 범위를 고정하기 위하여, 리뷰에 대해 다른 이용자들이 표시한 좋아요 수의 표준편차 값을 지수변환하여 사용하였다. 이러한 리뷰 인기도 값의 표준편차를 이용하여 리뷰가 판매 성과에 미치는 효과의 불확실성 값을 포착할 수 있기를 기대하였다. 리뷰 인기도의 효과의 크기를

포착할 수 있는 계수를 포함하였으며, 해당 계수는 각 제품별로 추정할 수 있도록 설정하였다. 시점별 제품 판매 여부와 리뷰 작성 여부는 각각 $d_{1,it}$ 와 $d_{2,it}$ 항이며, 두 변수 모두 1 값 일 때 실행되었으며, 0 값일때는 실행되지 않은 상태를 나타낸다.

3.2 제품 효과

제품의 특성 ($Product_{ik}$)의 효과는 제품별 계수인 γ_i 항으로 포착하고자 한다. 이 연구는 온라인에서 판매중인 도서를 대상 제품으로 분석하고자 하기 때문에, 제품의 특성으로 책과 관련된 책성이 포함되어 있다. 책 페이지수, 가격, 가격 할인 비율, 그리고 저자에 대한 고정효과 항을 분석 변수로써 벡터 항에 포함하였다. 책의 가격은 페이지가 많을수록 증가하는 경향이 있기 때문에, 페이지당 가격으로 변환한 값과 정규 가격을 모두 고려하였으나, 최종적으로 일반적인 정규가격을 사용하였다. 페이지당 가격은 페이지 수가 증가할수록 감소하는 경향이 있기 때문에, 일반적인 정규가격을 사용하였다. 따라서 정규가격의 효과에는 페이지에 따른 효과와 가격으로 인한 효용 효과가 혼합되어 있을 가능성이 있다. 하지만 페이지당 가격을 고려한 경우에는 정규가격으로 인한 효용의 효과를 포착하기 어렵다는 단점이 있기 때문에, 일반적인 정규가격을 분석 변수로 포함하였다. 도서는 저자의 인기도, 저자가 이전에 작업한 저작물의 유형 및 인기도 등이 도서 판매 성과에 중요한 요인이 될 수 있다. 따라서 이 연구와 같이 저자에 대한 고정효과 항을 포함함으로써 해당 효과를 통제하고 도서 자체의 효과를 보다 정확히 포착할 수 있다는 장점이 있다.

3.3 리뷰의 유형과 동적 스케일

이 연구는 제품에 대한 리뷰가 아이디어 공유 유형 또는 정보 공유 유형인지에 따라 성과에 미치는 차별적인 영향에 대해 설명하고자 한다. $Review_{ip}$ 변수는 각 리뷰의 내용에서 아이디어 공유와 관련된 내용의 비율이다. 각 리뷰의 텍스트를 새로운 아이디어를 제시 하는 내용과 제품 및 제작자 등과 관련된 정보를 제시하는 내용으로 구분하여 각 항목에 대한 점수를 부여하여 공유 내용에 대한 비율 점수를 계산하였다. 아이디어 공유 유형 또는 정보 공유 유형에 대한 점수는 3명의 독립된 평가자가 부여한 7점척도를 평균하여 두 유형에 대한 비율값을 산출하였다. 아이디어 유형 또는 정보 유형에 대한 기준이 되는 단어는 이전의 연구[9]를 참고하여 이 연구에 적용하였다.

만약 리뷰의 유형이 동일하더라도, 해당 리뷰가 최종 성과에 주는 효과는 차별적일 수 있다. 이전 연구에서는 리뷰 내용의 특성으로 문자수, 복잡성, 오타수, 공유 등을 포함하였으며, 이러한 변수는 리뷰 내용의 품질을 측정하는 근사 변수로 이용되었다[4][9][11][12]. 리뷰 내용의 품질은 리뷰에 대한 인기 및 다른 이용자들의 반응 및 최종적인 성과에 영향을 줄 가능성이 있다[12]. 따라서 모형 구축 시에 이러한 요인이 어떠한 영향을 줄 수 있을 지에 대해 설명할 수 있는 부분이 포함되어야 할 필요가 있다. 이 연구는 이러한 효과를 분석하기 위하여 리뷰 유형의 효과에 대한 동적 특성 스케일인 τ_{ikp} 항을 포함하여 리뷰 유형의 효과가 리뷰와 관련된 다른 특성에 의해 결정되는 정도를 설명하고자 하였다. 아래의 식(2)는 이러한 동적 스케일을 나타낸다.

$$\tau_{ikp} = \exp^{\theta_1 \cdot TextLen_{ip} + \theta_2 \cdot TextQual_{ip} + \theta_3 \cdot ReShare_{ip} + Time_{ikp} + \theta_{ikp}} \quad (2)$$

위 식의 각 항과 계수에 대한 정의는 아래와 같다.

$TextLen_{ip}$ = 제품 i 에 대한 리뷰 p 의 단어 수

$TextQual_{ip}$ = 제품 i 에 대한 리뷰 p 의 내용 중 오타 단어의 비율

$ReShare_{ip}$ = 제품 i 에 대한 리뷰 p 가 다른 이용자로 부터 제공유된 수

$Time_{ip}$ = 제품 i 에 대한 리뷰 p 가 작성된 시점에 대한 고정효과

ϱ_{ikp} = 정규분포를 따르는 오차항

위 식에서 리뷰 효과의 변화에 관한 동적 스케일은 지수 형태로 설정하였다. 리뷰의 효과가 0 이상의 비율 값을 가지는 범위 내에서 차이를 보일 것이라는 가정을 가지는 의미이다. 이는 리뷰 효과의 크기의 특성이 개별 리뷰의 효과와 다른 리뷰 사이에서 상대적인 차이를 설명 할 수 있다는 장점이 있다. 시간 효과($Time_{ip}$)는 리뷰를 작성한 월 단위에서 고정효과(fixed effect)를 포착하고자 하였다. 이때 시간효과는 리뷰를 작성한 시점으로부터 경과된 날짜의 수가 아닌 절대적 시점에 대한 명목변수로 포착한다.

3.4 이월 효과 함수

이 연구는 이월 효과를 계량 모형을 통하여 측정한 이전 연구[7][12][20]에 기반하여 시간에 따른 설명 요인이 가지는 효과의 감소 과정을 설명하기 위하여 이월 효과 함수를 제품 효과의 설명 변수에 포함하였다. 도서 카테고리에서 제품의 브랜드 가치에 해당하는 저자 정보를 이월효과의 설명 요인으로 포함하였는데, 저자의 이전 작품수, 저자의 해당 카테고리의 작품수, 그리고 저자의 이전 도서의 판매량의 합계를 포함하였으며, 고려하지 못한 요인은 정규분포를 따르는 오차항으로 기저선을 통제하였다. 이월효과 함수는 다음의 식(3)과 같다.

$$\delta_{im} = \exp^{\rho_1 \cdot NumProduct_{im} + \rho_2 \cdot NumCategory_{im} + \rho_3 \cdot TotalSales_{im} + \varpi_{im}} \quad (3)$$

위 식에서 다루고 있는 변수 및 계수에 대한 정의는 아래와 같다.

$NumProduct_{im}$ = 시점 m 에서 제품 i 의 제작자의 총 제품수

$Category_{im}$ = 시점 m 에서 제품 i 와 같은 카테고리의 제품수

$TotalSales_{im}$ = 시점 m 에서 제품 i 의 제작자의 누적 판매량

ϖ_{im} = 정규분포를 따르는 오차항

3.5 데이터

이 연구는 제품에 대한 리뷰의 유형에 따라서 제품의 판매 성과에 미치는 차별적인 영향을 장기간에 걸쳐 추적 및 분석할 수 있는 데이터 구조와 모형을 제시하여 실증적으로 분석하고자 한다. 이러한 연구 목적과 실증 분석 과정에 적합한 데이터를 선정하였다. 주요 분석 대상은 온라인 서점의 도서별 판매 부수와 해당 도서에 대해 일반 이용자가 작성한 리뷰이다. 온라인 서점은 터키 최대의 온라인 서점중 하나인 Kitapyurdu로 선정하였다. 해당 온라인 서점에서는 도서별 판매수가 실시간으로

제공된다는 장점이 있다. 베스트셀러 목록 중 상위 100개의 도서를 대상 제품으로 선정하였다. 베스트셀러 목록 수집 시점은 2022년 1월이며, 전체 판매 데이터 및 리뷰 데이터의 수집 기간은 2022년 1월부터 3월까지의 3개월 동안이며 일 단위에서 수집하였다. 도서 데이터에는 책 목록을 기반으로 책의 제목, 페이지수, 정규가격, 할인비율, 기타 프로모션 여부, 배송 기간 등의 변수를 수집하였다. 도서의 저자 정보는 온라인 서점에서 제공한 데이터를 기반으로, 온라인 검색을 추가로 실시하였다. 저자 정보에는 저자의 다른 도서 목록, 다른 도서의 판매량 등의 변수를 수집하였다.

리뷰 데이터는 트위터를 대상으로 수집하였다. 트윗 내용에서 책 제목 또는 책에 대한 링크가 있는 경우를 검색하여 텍스트를 수집하였다. 데이터 수집 도구는 트위터에서 제공하는 API를 이용한 자체제작 프로그램이다. 각 트윗 내용 이외에도 작성 날짜, 트윗 작성자의 팔로워수, 해당 트윗에 대한 좋아요, 리트윗(e.g., 재공유) 횟수를 수집하였다.

아래의 [표 2]는 수집한 데이터에 대한 기초 통계이다. 총 분석 대상 도서 100개에 대한 데이터를 3개월 간 수집하여 8972개의 표본을 수집하였다. 데이터의 일반적인 특성으로, 판매 성과는 평균적으로 약 5만여 개이며, 약 40%에 가까운 할인율을 보임으로써, 해외 도서 판매점의 할인 비율이 비교적 높은 것을 알 수 있다. 제작자(도서의 저자)별 특성을 살펴보면, 평균적으로 4.6권의 저술을 가지고 있으며, 이중 소설의 비중이 절반 이상임을 알 수 있다. 트윗을 통한 도서 리뷰의 경우, 평균적으로 25개의 단어를 작성하였으며, 이중 오타의 비율이 약 5.93%인 것을 알 수 있다. 리뷰 작성자의 평균적인 팔로워 수는 605명이었다.

[표 2] 기초통계

[Table 2] Summary Statistics (N=8972)

| Category | Variable | Mean | STD. |
|----------|--------------------|----------|-----------|
| 종속 변수 | 판매 성과 | 53857.25 | 77484.11 |
| 제품 성과 모형 | 리뷰 유형%(1=아이디어 공유형) | 14.10 | 36.07 |
| 제품 효과 함수 | 페이지수 | 229.85 | 115.41 |
| | 정규가격(천원) | 10.70 | 9.31 |
| | 할인율(%) | 38.77 | 7.98 |
| | 기타 프로모션 비율 | 65.80 | 24.03 |
| 이월효과 함수 | 총 제품수 | 4.06 | 11.15 |
| | 카테고리의 제품수(소설) | 2.96 | 5.81 |
| | 누적판매량 | 95018.40 | 193609.60 |
| 동적 스케일 | 단어 수 | 25.13 | 20.18 |
| | 오타 단어의 비율 (%) | 5.93 | 5.40 |
| | 재공유 수 | 4.52 | 9.32 |
| | 팔로워 수 | 605.26 | 1345.74 |

4. 분석 결과

이 연구는 정교하게 설계된 여러 파라미터의 동태적 상태에 대한 추정을 할 수 있는 MSM (method of simulated moment) 방법을 이용하여 모형을 추정하였다[20]. 분석과정은 SAS PROC model을 이용하여 초기값을 도출한 이후에, 해당 초기 값을 이용하여

Matlab에서 작성한 코드를 이용하여 분석하였다. 식 (1)의 동태적 성과 모형은 여러 계수들의 상호 연쇄적인 관계를 다루고 있으므로, MSM 방법이 적합하다고 할 수 있다. 이 연구는 상당히 많은 계수를 추정할 필요가 있으므로, 일반 적인 시계열 모형이나 GLM(Generalized Linear Model)을 이용한 방법으로는 추정이 어려울 수 있다. 동태적 성과 모형을 추정하는 과정에서 리뷰 유형의 동적 스케일 및 이월 함수를 동시에 추정하였다. 모형 추정 결과는 다음의 [표 3]이다.

[표 3] 모형 추정 결과

[Table 3] Estimateion Result

| Category | Variable | Estimate | S.E. |
|----------|--------------------|----------|------|
| 제품 성과 모형 | 리뷰 유형 (아이디어 공유형=1) | 0.01* | 0.00 |
| 제품 효과 함수 | 페이지수 | 0.01 | 0.01 |
| | 정규가격 | -0.24* | 0.10 |
| | 할인율 | 0.11* | 0.05 |
| | 프로모션여부 | 0.08 | 0.07 |
| 동적 스케일 | 단어 수 | 0.03* | 0.01 |
| | 오타 단어의 비율 | 0.00 | 0.01 |
| | 재공유 수 | 0.02* | 0.01 |
| | 팔로워 수 | 0.01* | 0.00 |
| | 시간통제: 1월 | 0.00 | 0.01 |
| | 시간통제: 2월 | 0.01 | 0.01 |
| 이월 효과 함수 | 총 제품수 | 0.05* | 0.02 |
| | 카테고리의 제품수(소셜) | 0.04 | 0.05 |
| | 누적판매량 | 0.09* | 0.02 |

Note: * p-value < 0.01

분석 결과에 따르면 제품 성과 모형에서 리뷰 유형의 계수가 0.01로 유의미 하게 추정된것으로 나타났으며, 이는 리뷰의 내용이 아이디어 공유에 가까울 수록 제품 성과가 더 높은 것을 의미한다. 제품 효과 함수에서는 정규가격이 -0.24, 그리고 할인율이 0.11로 유의미 하게 나타났으나, 페이지수 및 프로모션 여부는 비유의한 값으로 도출되었다. 따라서, 제품의 특성이 판매 성과에 미치는 효과를 설명하는데 있어서, 가격이 낮고 할인율이 높을수록 제품 성과가 더 클 수 있음을 예상 할 수 있다. 반면 페이지수 및 프로모션 여부는 비유의한 것으로 나타났는데, 책의 페이지 수는 책의 무게와 비례하는 관계에 있는 것으로 예상할 수 있으므로, 책의 가격과 다소 관계가 있을 것으로 예상할 수 있다. 다만 책의 가격에 비해 무게 및 페이지수 등은 제품의 가치에 상대적으로 적은 영향을 줄 수 있다는 점을 추측해 볼 수 있다.

아이디어 공유형 리뷰의 장기간에 걸친 누적 효과를 설명하기 위한 동적 스케일의 설명 요인에서는 단어 수의 계수가 0.03, 재공유 수가 0.02, 그리고 팔로워 수가 0.01로 유의미한 값으로 도출되었다. 따라서, 길게 작성한 리뷰, 재공유를 많이 받은 리뷰, 그리고 팔로워 수가 많은 트위터의 리뷰가 장기적 효과가 클 수 있음 알 수 있다. 또한 동적 스케일은 리뷰의 효과를 설명하는 함수이며, 지수형태로 설정되어 있다. 따라서

길이, 재공유, 팔로워가 증가할수록 리뷰의 효과는 지수적으로 급격히 증가함을 예상할 수 있다.

마지막으로 이월 효과의 경우 저자가 과거에 출판한 도서가 많고 이전의 누적 판매량이 많을수록 이월 효과가 더 크게 나타 났다. 즉 브랜드 파워가 큰 저자의 작품일수록 다른 이용자의 리뷰에서 아이디어 공유로 인한 긍정적 판매 성과의 효과가 더 장기간에 걸쳐 이전될 수 있음을 암시한다. 온라인에서 판매되는 도서제품의 성과는 작가의 이전 활동이 현재 제품에 전이될 뿐만 아니라, 작가의 이전 활동이 많을수록 리뷰 및 제품 특성의 효과가 장기적으로 지속되는 것을 알 수 있다. 이러한 과정은 소비자가 제품을 판단하는 근거로 활용하는 제품 리뷰 및 제품 특성 정보를 판단하는 데있어서 작가의 이전 활동을 근거로 할 수 있음을 추측할 수 있다.

5. 결론

이 연구는 도서 제품의 장기적인 판매 성과의 변화를 결정하는 요인을 설명하기 위하여, 제품에 대한 리뷰의 내용을 분석하였다. 리뷰를 아이디어 공유형과 정보 공유형으로 구분하여, 각 유형의 리뷰가 제품 판매 성과에 미치는 차별적인 영향에 대해 설명하였다. 실증 분석 결과는 다음과 같이 요약할 수 있다. 첫째, 제품 성과 모형에 포함된 주요 변수인 리뷰 유형은 아이디어 공유형일때 정보 공유에 비해 제품 성과에 더 큰 영향을 주는 것으로 나타났다. 아이디어 공유형 리뷰의 작성 과정에 더 많은 정교성과 노력이 포함되었다는 사실을 다른 이용자가 인지하게 됨으로써 제품 성과에 더 큰 영향을 줄 수 있음을 암시하는 결과라고 할 수 있을 것이다. 이러한 결과는 아이디어 공유형이 이용자들의 관심과 자발적인 참여 효과를 증가시킴으로써 제품에 대한 몰입과 재이용에 대한 성과를 높일 수 있다는 이전 연구의 내용과 일치하는 결과를 보였다[9][13]. 둘째, 동적 스케일에 포함된 단어수, 재공유수, 팔로워수 등이 유의미한 효과가 있는 것으로 나타났는데, 이는 리뷰 효과가 이러한 요인에 의해 상승과 감소가 결정 될수 있음을 나타낸다. 따라서, 리뷰의 유형이 동일하더라도, 해당 하부 요인으로 인하여 리뷰의 효과가 달라 질 수 있음을 나타낸다. 저자의 누적된 브랜드 성과가 높을수록 아이디어 공유 리뷰의 효과가 더 장기적으로 지속될 수 있음을 확인할 수 있다. 셋째, 이월효과가 유의미하게 나타남으로써 제품 함수 및 리뷰 함수의 원인 변수의 효과가 장기적으로 유지되는 것을 알 수 있다. 이외에도 제품 효과 함수에서는 정규가격과 할인율과 같은 요인이 제품 성과에 영향을 줄 수 있는 것으로 나타났다.

이 연구의 이론적 시사점은 다음과 같다. 첫째, 리뷰의 질적 차이에 관한 연구를 실제 제품 성과 분야로 확장하였다. 이전 연구[7][9][18]에서는 리뷰의 질적 차이가 브랜드 가치 등에 미치는 과정을 설명하였으나, 이 연구는 실제 판매성과를 분석함으로써 이론 분야의 연구 성과에 기여하였다고 할 수 있을 것이다. 둘째, 이용자생산컨텐츠 연구 분야에서 리뷰의 질적 차이를 측정하여 반영할 수 있는 실증적 모형을 제시하였다. 이 연구는 장기적 브랜드 성과 측정 모형을 이용자생산컨텐츠로 확장하여 적용하였으며, 이에 대한 모형 구축과정에 대한 제시를 하였고, 이후 다른 연구의 발판이 될것으로 기대한다. 마지막으로, 외부 소셜 네트워크가 제품 성과에 미치는 효과에 대한 연구를 확장하였다. 이 연구는 온라인 서점과 트위터 데이터를 결합하여 분석함으로써, 외부 네트워크 효과를 검증하였다는 의의가 있다.

이 연구는 다음과 같은 실무적 시사점을 제공한다. 이 연구는 온라인 상점에서 제품을

구매하고 리뷰를 작성하는 이용자들의 활동 데이터를 수집하고 실제 판매 성과를 분석하였다는 점에서 온라인 서점 및 온라인 소셜 네트워크를 관리하는 기업 및 조직에게 전략적 가이드 라인을 제시할 수 있다. 현재 온라인 상점에서는 리뷰의 양과 빈도에 대한 관심이 높지만, 리뷰 내용의 유형에 대한 관심은 상대적으로 낮은 편이라고 할 수 있다. 이 연구에서와 같이 리뷰의 내용을 구분할 수 있는 체계를 구축하여, 온라인 미디어 및 소매점에서 전략적으로 활용할 수 있을 것으로 기대한다. 또한 외부 소셜 네트워크의 온라인 리뷰는 장기적인 효과가 있으므로, 이 연구의 분석 결과와 같이 이용자생산콘텐츠 제작자의 특성과 제품 특성을 모두 고려하여, 제품 성과에 미치는 효과의 우선순위를 선정하여 관리 할 필요가 있음을 시사한다.

이 연구가 리뷰의 유형을 구분하여 실제 온라인 판매 성과에 미치는 효과를 분석할 수 있는 체계를 제시하였고 실증적 결과를 도출하였음에도 불구하고, 향후 추가적인 연구가 지속적으로 이루어질 필요가 있다. 이 연구에서는 이전 연구에 기반하여 리뷰의 유형을 두 가지로 구분하였으나 소비자의 감정 및 경험 등과 같은 다양한 내용을 구분할 수 있는 기준을 적용할 수 있는 추가 연구가 필요할 것이다. 리뷰는 이용자들의 개인적인 감정 및 경험에 기반한 경우가 빈번하기 때문에 이에 대한 고려가 필요할 것이다[4][5]. 온라인 서점에 관한 이전 연구에 따르면, 롱테일 제품 군의 경우 상위 제품에 대한 소비자의 선택 프로세스가 다르게 작동할 수 있으므로, 하위 인기 제품에 대한 추가적인 분석과 설명이 필요할 것이다. 제품의 특성으로 고려한 다양한 세부적인 요소를 모형과 데이터 구축과정에 세심하게 고려할 필요가 있을 것이다.

6. 감사의 글

이 논문은 충남대학교 2018년 자체연구과제로 수행되었음.

References

- [1] https://www.motie.go.kr/motie/ne/presse/press2/bbs/bbsView.do?bbs_seq_n=166599&bbs_cd_n=81¤tPage=1&search_key_n=&cate_n=&dept_v=&search_val_v=, Nov 30 (2022)
- [2] <https://news.einfomax.co.kr/news/articleView.html?idxno=4252729>, Feb 01 (2023)
- [3] J. Chevalier, D. Mayzlin, The Effect of Word of Mouth on Sales: Online Book Reviews, *Journal of Marketing Research*, (2006), Vol.43, No.3, pp.345-354.
- [4] J. Berger, K. L. Milkman, What Makes Online Content Viral?, *Journal of Marketing Research*, (2012), Vol.49, No.2, pp.192-205.
- [5] J. H. Chung, Y. C. Kim, An Analysis of WOM Effects on the Consumer Product Choice by Using a Hierarchical Bayesian Probit Model, *Journal of Korean Marketing Association*, (2004), Vol.19, No.3, pp.1-20.
- [6] S. J. Lee, I.C. Choi, Effect of Rating Valence and Object Types (Artwork vs. Artifact) on Consumer Purchase Intention in an Online Shopping Context, (2022), Vol.37, No.1, pp.1-19.
- [7] S. Melumad, J. J. Inman, M. T. Pham, Selectively Emotional: How Smartphone Use Changes User-Generated Content, *Journal of Marketing Research*, (2019), Vol.56, No.2, pp.259-275.
- [8] C. Kim, F. J. Costello, K. C. Lee, The Unobserved Heterogeneous Influence of Gamification and Novelty-Seeking Traits on Consumers' Repurchase Intention in the Omnichannel Retailing, *Frontiers in Psychology*, (2020), Vol.11, pp.1664-1685.

- [9] U. Narang, M. S. Yadav, A. Rindfleisch, The idea advantage: How content sharing strategies impact engagement in online learning platforms, *Journal of Marketing Research*, (2022), Vol.59, No.1, pp.61-78.
- [10] S. L. Jeon, G. H. Kim, The Influence of Two-sided WOM on Consumers' Brand Attitudes in an Online Context, *Journal of Korean Marketing Association*, (2012), Vol.27, No.1, pp.45-66.
- [11] G. Packard, J. Berger, How Language Shapes Word of Mouth's Impact, *Journal of Marketing Research*, (2017), Vol.54, No.4, pp.572-88.
- [12] A. Ghose, P. G. Ipeirotis, B. Li, Designing ranking systems for hotels on travel search engines by mining user-generated and crowdsourced content, *Marketing Science*, (2012), Vol.31, No.3, pp.493-520.
- [13] B. H. Thompson, G. H. Thompson, Confronting Diversity Issues in the Classroom with Strategies to Improve Satisfaction and Retention of International Students, *Journal of Education for Business*, (1996), Vol.72, No.1, pp.53-55.
- [14] E. Brynjolfsson, Y. Hu, M. D. Smith, Consumer surplus in the digital economy: Estimating the value of increased product variety at online booksellers, *Management Science*, (2003), Vol.49, No.11, pp.1580-1596.
- [15] J. H. Seok, Y. S. Lee, B. D. Kim, Research on the Interrelationship Among Volume and Valence of eWOM and the Number of Visitors of a Museum, (2020), Vol.35, No.3, pp.1-23.
- [16] J. Berger, Arousal increases social transmission of information, *Psychological science*, (2011), Vol.22, No.7, pp.891-893.
- [17] A. Ghose, P. G. Ipeirotis, B. Li, Modeling consumer footprints on search engines: An interplay with social media. *Management Science*, (2019), Vol.65, No.3, pp.1363-1385.
- [18] E. R. Smith, J. DeCoster, Dual-Process Models in Social and Cognitive Psychology: Conceptual Integration and Links to Underlying Memory Systems, *Personality and Social Psychology Review*, (2000), Vol.4, No.2, pp.108-131.
- [19] J. Son, A Study on the Dynamics of User Contribution Performance by the Pre- and Post-Producer Participation in User-Generated Content Platform, *Journal of Korean Marketing Association*, (2016), Vol.31, No.2, pp.31-55.
- [20] M. Trusov, R. E. Bucklin, K. Pauwels, Effects of word-of-mouth versus traditional marketing: findings from an internet social networking site, *Journal of Marketing*, (2009), Vol.73, No.5, pp.90-102.
- [21] L. Luo, X. Chen, J. Han, C. W. Park, Dilution and enhancement of celebrity brands through sequential movie releases, *Journal of Marketing Research*, (2010), Vol.47, No.6, pp.1114-1128.
- [22] J. Son, J. Lee, The Models for the Dynamic Brand Value of Content Producers in the Online Platform, *Journal of Convergence for Information Technology*, (2022), Vol.12, No.5, pp.92-99.
- [23] S. C. Ahn, P. Schmidt, Efficient estimation of dynamic panel data models: Alternative assumptions and simplified estimation, *Journal of Econometrics*, (1997), Vol.76, No.1-2, pp.309-321.