KBR 제30권 제2호 2026년 5월 http://dx.doi.org/10.17287/kbr.2026.30.2.237

Emotional Attributes of Films and Their Influence on Diffusion Patterns: An Exploratory Analysis Using the Bass Model*

Jihwan Lee**
Youseok Lee***
Sang-Hoon Kim****

This study explores how the emotional attributes of films, inferred from online review texts, influence their diffusion patterns in the Korean movie market. Using deep learning‑based emotion classification (KcELECTRA) on 182,987 audience reviews, 39 emotion variables were extracted and reduced via factor analysis to nine dimensions. The estimated Bass model parameters—innovation (p), imitation (q), and market potential (m)—served as dependent variables in regression analyses. Drawing on Affect-as-Information and Arousal-driven Sharing theories, this study elucidates the psychological mechanisms linking specific emotions to market behaviors. Results show that *Tension* and *Irritation* exert opposite effects on innovation and market potential, while *Tension* also enhances imitation, suggesting that films evoking suspense or unpredictability attract word-of-mouth‑driven audiences and achieve stronger long-term performance. Critically, these emotional effects remain significant even after controlling for structural factors such as genre and screen availability. These findings provide new empirical evidence linking emotion-based textual signals to film diffusion dynamics and offer practical insights for producers and marketers to balance emotional tone and audience engagement strategies in the evolving, post-pandemic film industry.

Key Words: Affect-as-information, Arousal-driven sharing, Bass Diffusion Model, Emotion Analysis, Movie Industry, Natural Language Processing, Online Reviews, Social Sharing of Emotion

Ⅰ. Introduction

The global movie industry underwent significant upheaval during the COVID-19 pandemic, and the Korean film market was no exception (Korean Film Council, 2024). Prior to the pandemic, Korea stood out for its high theater attendance rate, with an average of 4.37 visits per capita in 2019, and a

Received: 2025. 06. 04. First Revision Received: 2025. 11. 16, 2nd Revision Received: 2026. 01. 23, 3rd Revision Received: 2026. 03. 04. Accepted: 2026. 03. 10.
* This study was supported by the Institute of Management Research at Seoul National University.
** Samjong KPMG(hwankhai@snu.ac.kr), First Author
*** Associate Professor, Department of Business Administration, Dongguk Business School, Dongguk University-Seoul(yslee2@dongguk.edu)
**** Professor, Graduate School of Business, Seoul National University(profkim@snu.ac.kr), Corresponding Author

domestic film market size of approximately 1.91 trillion KRW. This robust market was characterized by a balanced local film share of around 51%. However, the pandemic caused a marked downturn, reducing the 2021 market size to below half of its 2019 level (Korean Film Council, 2024). Although partial recovery is underway, theater attendance and box-office revenues remain below pre-pandemic levels. Concurrently, the growing influence of streaming services and evolving consumer preferences is altering traditional movie consumption patterns, further complicating the market environment. Therefore, it is necessary to conduct an in-depth exploration of the diffusion patterns of films under such circumstances.

This study aims to exploratorily examine the influence of films' emotional attributes on their diffusion patterns. The emotional attributes of each film are inferred by text- mining review data written by actual moviegoers. Furthermore, the diffusion patterns are estimated by applying the Bass model to daily audience data, through which the study empirically investigates how the emotional attributes of films affect the parameters derived from the model. The Bass diffusion model (Bass, 1969) provides a foundational framework for understanding the adoption and spread of new products, including movies. This model posits that consumers can be categorized into two groups: innovators, who are influenced by external factors such as advertising and media, and imitators, who are influenced by the internal factor of word-of-mouth from prior adopters. The interplay between these two groups drives the overall diffusion pattern. This study employs a Generalized Bass Model (GBM) to account for the inherent seasonality in movie attendance, such as increased viewership on weekends and holidays (Lee et al., 2017).

Although numerous previous studies have attempted to identify factors influencing the diffusion patterns of films, few have directly measured the emotional attributes of films themselves. Moreover, while many studies have examined the relationship between audience review data and box office performance, there has been little research that extracts films' emotional attributes from review data as the basis for analysis. While prior research has established that review volume and valence affect sales (Liu, 2006), the specific psychological mechanisms by which distinct discrete emotions drive diffusion parameters remain underexplored. For instance, why do some negative emotions like irritation accelerate early adoption while others suppress it? To address this gap, we ground our investigation in three established psychological frameworks.

First, the Affect-as-Information Theory (Schwarz & Clore, 1983; 1996) posits that individuals frequently rely on their current affective state as a heuristic source of information for decision-making ("How do I feel about this?"). This mechanism is particularly relevant for innovators in the Bass model (p), who make independent adoption decisions. We hypothesize that high-arousal emotions, even if negative (e.g., irritation), may serve as heuristic cues signaling the importance or controversy of a film, thereby acting as a catalyst for early adoption among innovators who value novelty.

Second, to explain the imitation effect (q), we draw on the Arousal-Driven Sharing Theory (Berger, 2011) and the Social Sharing of Emotion framework (Rimé, 2009). These theories suggest that physiological arousal is a key driver of social transmission. High-arousal emotions, such as tension or anger, create a state of psychological disequilibrium that individuals seek to resolve through communication and sharing. Thus, we expect that films evoking high-arousal emotions will exhibit a higher imitation coefficient (q), reflecting accelerated word-of-mouth diffusion, compared to films evoking low-arousal emotions like sadness or contentment.

Online reviews have become increasingly influential in shaping consumer information acquisition and decision-making, especially for experiential products like movies, whose quality is difficult to assess before consumption (Chevalier & Mayzlin, 2006; Zhu & Zhang, 2010). Consumers actively seek out online reviews to gather information, form expectations, and reduce uncertainty about a movie's quality, ultimately impacting box office performance (Liu, 2006). While early research focused on the impact of review volume and valence (e.g., positive vs. negative), recent studies have recognized the importance of analyzing the emotional content embedded within review texts (Ludwig et al., 2013). Ullah et al. (2025) empirically examined the effect of the proportions of positive and negative emotion words extracted from movie review texts on box office revenue. However, the present study is based on the premise that the emotions expressed in online reviews are elicited by the viewing experience itself and, therefore, can serve as indicators of a film's emotional attributes. Building on this assumption, the study seeks to examine how these emotional attributes influence the diffusion of films.

Recent advances in text mining and natural language processing have enabled researchers to extract and analyze emotional content from large-scale text data (Kim et al., 2018). For instance, Berger et al. (2023) leveraged natural language processing to analyze over 600,000 reading sessions from more than 35,000 pieces of content, demonstrating that text which was easier to process (more readable, syntactically simple, and using familiar or concrete words) helped sustain reader attention. Jang et al. (2022) employed an emotion vocabulary dictionary to annotate a movie review corpus with nine discrete emotions—joy, sadness, fear, anger, disgust, surprise, interest, boredom, and pain—allowing for nuanced emotion classification. Mokryn et al. (2020) empirically demonstrated that emotions extracted from movie review texts can be regarded as proxies for a film's intrinsic attributes and can predict box office performance. However, the operationalized emotional variables were limited to five—four positive emotions and the aggregate of several negative emotions.

By analyzing various emotional attributes from online reviews, the current study explores their differential impacts on the movie diffusion process. Conventional sentiment analysis approaches often rely on the unrealistic assumption that most consumers decide to watch a film after reading other consumers' reviews. In particular, the present study employs the Bass model to estimate diffusion patterns, wherein innovators are defined as consumers who make viewing decisions independently of others' word-of-mouth influence. Therefore, the assumption that the content of review texts directly determines the overall diffusion of a film does not align with reality. In contrast, this study adopts a distinctive approach by text-mining review data to infer the emotional attributes of films, thereby demonstrating that the proportions of innovators and imitators may vary depending on the emotional characteristics embedded in each film.

The results reveal that specific emotional attributes of films, particularly tension-related emotions, significantly shape diffusion dynamics as captured by the Bass model. This study contributes to the emerging literature by empirically linking emotion-based textual signals to market diffusion patterns, offering new insights into how film content characteristics predict commercial outcomes in the Korean film industry.

Ⅱ. Methods

2.1 Sample

Initial data comprised daily box office records (January 1, 2022 – November 1, 2024) from the Korean Film Council's KOBIS, supplemented by public holiday data. Preprocessing involved several exclusions, including films released before April 17, 2022 (the end of social distancing), box office data beyond 100 days post-release, films released after September 13, 2024 (to ensure sufficient observation), pre-release audience data, re-releases, and films with fewer than 100,000 cumulative viewers. This yielded 219 films. As online reviews were sourced from CGV, a major cinema chain, four films not released on this platform were further excluded, resulting in a candidate pool of 215 films (Choe & Kang, 2017; Kwon & Kim, 2021). Final sample determination based on Bass model estimation is detailed in Section 3.1.

2.2 Variables

2.2.1 Dependent Variables

Movie diffusion was modeled using a Generalized Bass Model (Bass et al., 1994), capturing innovation effect (p), imitation effect (q), and market potential (m). The coefficient of innovation (p) represents the likelihood that an individual will adopt a product independently of others—typically driven by marketing efforts or mass media. Within our theoretical framework, this parameter captures the behavior of consumers who rely on external signals and heuristic affective cues (Affect-as-Information) to make early adoption decisions. The coefficient of imitation (q) reflects the tendency for adoption to occur as a result of interpersonal communication and social influence, indicating the strength of diffusion through word-of-mouth. We conceptualize this parameter as being driven by the social sharing imperative of high-arousal emotions (Arousal-Driven Sharing). Finally, the market potential (m) denotes the total number of potential adopters in the population, corresponding to the overall market size for the product. The model, incorporating day-of-week and holiday effects (Lee et al., 2017), is expressed as

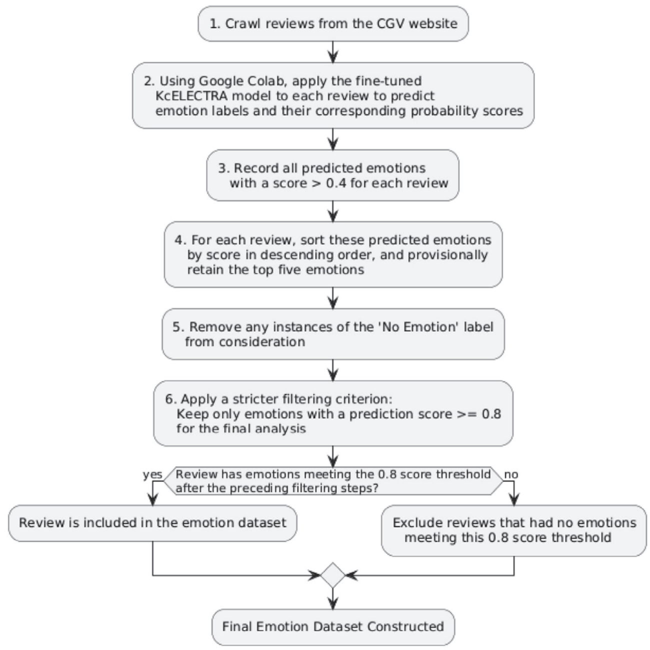
$$audiCnt_t = \left(p + q \cdot \frac{\sum_{i=1}^{t-1} audiCnt_i}{m}\right) \cdot \left(m - \sum_{i=1}^{t-1} audiCnt_i\right) \cdot \left(1 + \beta_1 \cdot fri + \beta_2 \cdot (sat + sun + hol)\right) \quad (1)$$

where $\boldsymbol{audiCnt_t}$ is the daily audience at time $\boldsymbol{t}$; $\boldsymbol{\Sigma_{i=1}^{t-1} audiCnt_i}$ is the cumulative audience up to time $\boldsymbol{t-1}$; $\boldsymbol{fri}$, $\boldsymbol{sat}$, $\boldsymbol{sun}$, and hol are dummy variables for Friday, Saturday, Sunday, and public holidays, respectively; and β_1 and β_2 are coefficients capturing the respective day effects. The parameters p, q, and m are estimated for each film using daily audience data. The specific Non-Linear Least Squares (NLS) estimation procedure and results are detailed in Section 3.1. These estimated p, q, and m values for each film serve as the dependent variables in the subsequent regression analyses.

2.2.2 Independent Variables

Emotion ratios from CGV online reviews served as independent variables. The initial aim was to collect the first 1,000 reviews per movie post-release. Emotion analysis utilized a KcELECTRA model (Lee, 2021) fine-tuned with the KOTE dataset (Jeon et al., 2022), which identifies 43 Korean emotion categories.

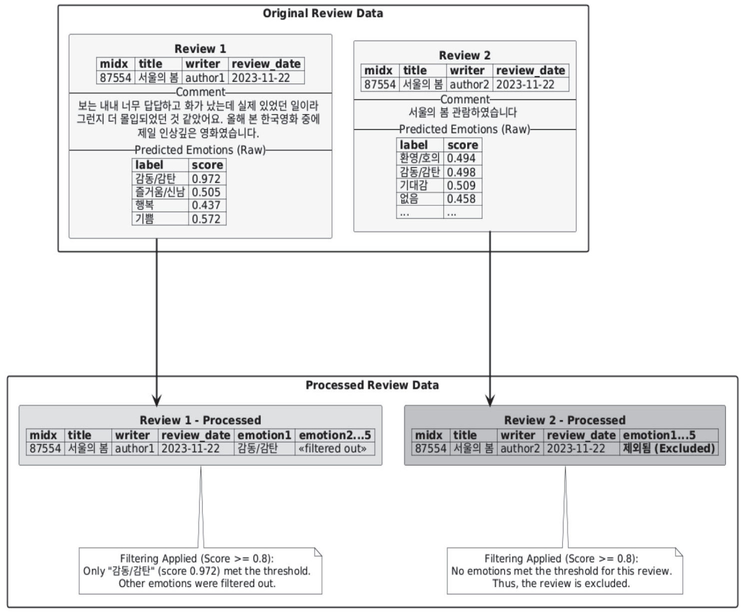
The process of constructing the emotion dataset for each movie involved several steps, visualized in 〈Figure 1〉. First, reviews were crawled from the CGV website. Second, using Google Colab, the fine-tuned KcELECTRA model was applied to each review to predict emotion labels and their corresponding probability scores. The model can assign multiple emotion labels to a single review, and these scores represent in-



〈Figure 1〉 Emotion Data Preprocessing Process

dependent probabilities for each label, thus not summing to one. Initially, all predicted emotions with a score greater than .4 were recorded for each review. Third, for each review, these predicted emotions were sorted by score in descending order, and the top five emotions were provisionally retained. Fourth, any instances of the 'No Emotion' label were removed from consideration. Fifth, a stricter filtering criterion was applied: only emotions with a prediction score of .8 or higher were kept for the final analysis. This threshold was determined based on a qualitative assessment of prediction accuracy against review content to ensure a high confidence level in the identified emotions. Sixth, reviews that had no emotions meeting this .8 score threshold after the preceding filtering steps were excluded from the emotion dataset.

This rigorous filtering process resulted in a final baseline emotion dataset comprising 182,987 reviews across the analyzed films. By aggregating such a large volume of reviews per film, we leverage the Law of Large Numbers to mitigate the influence of individual-level noise (e.g., a viewer's personal mood or viewing context). While individual reviews may be influenced by external factors, the aggregated emotion ratios effectively capture the common sentiment driven by the film's intrinsic attributes, thus ensuring ecological validity. The number of reviews per movie varied, with a min-



〈Figure 2〉 Example of Emotion Extraction from a Review

imum of 411 and a maximum of 994. An example of how emotions are extracted from a review text and their associated scores can be seen in 〈Figure 2〉. After applying these filters, 39 distinct emotion categories (excluding 'No Emotion' and four other emotions that either never appeared or never met the .8 score threshold) were present in the dataset.

For each of the 215 films and for each of the 39 emotions, an emotion ratio ($R_{e,m}$) was calculated as

$$R_{e,m} = \frac{N_{e,m}}{K_m} \tag{2}$$

where $N_{e,m}$ is the number of reviews for movie m that express emotion e (with a score $\geq$ .8), and K_m is the total number of valid reviews analyzed for movie m. These emotion ratios served as the primary independent variables. Given the large number of emotion variables, factor analysis was employed to reduce dimensionality, as detailed in Section 4.2. The resulting factor scores were then used in regression models to examine their relationship with the Bass diffusion parameters.

2.2.3 Control Variables

To examine the relationship between emotions in online reviews and film diffusion patterns more rigorously, we incorporated several control variables. Following Bass (1969), the innovation coefficient p is affected by exogenous factors such as marketing communication and product availability, while the imitation coefficient q is influenced by endogenous factors such as word of mouth. Because marketing expenditure data are unavailable in the Korean film market, we used the number of poster images and trailer videos on each film's Naver movie page as proxies (Lee et al., 2021). Product availability was measured by the number of nationwide screens at release (Lee et al., 2017).

Word of mouth was captured through its volume and valence, operationalized as the total number and average score of Netizen ratings on Naver (Dellarocas et al., 2007). We also controlled for seasonality, given that low-involvement audiences, who are more responsive to word of mouth, are more prevalent during peak seasons. Dummy variables indicated whether the first two weeks of release overlapped with school vacation months (January–February, July–August) or major holidays (Seollal, Chuseok) (Lee et al., 2021). Finally, we included 19 dummy variables representing film genres (e.g., Action, Thriller, Drama) derived from KOBIS data. This is a critical methodological step to disentangle the influence of structural factors from emotional attributes. By controlling for genre, our analysis isolates the marginal effect of a film's specific emotional execution beyond the baseline expectations of its genre.

Ⅲ. Results

3.1 Bass Diffusion Model Estimation

Bass parameters (p, q, m) were estimated using NLS with constraints (p, q $\geq$ .00001, and m 〉 observed maximum audience). 〈Table 1〉 shows descriptive statistics for the estimated parameters from convergent models. From the 215 candidate films, 6 films with non-significant p and q coefficients were excluded, resulting in a final sample of 209 films for analysis.

3.2 Factor Analysis of Emotional Variables

Factor analysis (Principal Component Analysis with Varimax rotation; loadings 〈 .3 suppressed) on the 39 emotion ratios yielded nine factors. 〈Table 2〉 shows the rotated component matrix. Factor scores were computed for each film on these nine dimensions

〈Table 1〉 Descriptive Statistics of Bass Model Parameter Estimates (N=209)

Parameter	Mean	S.D.	Min	Max
p	0.0719	0.0409	0.0071	0.1948
q	0.0299	0.0356	0.0001	0.1642
m	1,436,695	2,401,305	99,084	13,185,081

〈Table 2〉 Factor Analysis Results for 39 Emotions (Rotated Component Matrix)

	1	2	3	4	5	6	7	8	9
Dissatisfaction	0.961								
Irritation	0.943								
Preposterous	0.933								
Boredom	0.930								
Disappointment	0.907								
Pathetic	0.833								
Admiration	-0.721				0.315				
Embarrassment	0.647			0.547					
Joy	-0.572	-0.302			-0.488	-0.391			
Distrust		0.882							
Despair		0.875							
Anger		0.868							
Resolute		0.825							
Sadness		0.732				0.497			
Compassion		0.726				0.485			
Contempt		0.722	0.525						
Fed up		0.602							
Disgust			0.878						
Sorrow			0.803						
Shock			0.718						
Anxiety		0.383		0.754					
Fear			0.584	0.687					
Surprise			0.410	0.620	0.312			-0.367	
Reluctant				0.563					
Comfort					-0.703				
Expectancy					0.680				
Attracted					-0.581				0.321
Happiness	-0.500				-0.571				
Exhaustion						0.822			
Realization						0.738			
Excitement	-0.488	-0.339				-0.508	-0.353		
Relief							0.811		
Respect							0.804		
Gratitude							0.710	0.561	
Care								0.777	
Interest					0.352			-0.510	
Welcome	-0.367			-0.311				0.454	
Pride									0.781
Arrogancy									0.361

(labeled Factor 1 to Factor 9) for use in regression.

Based on the factor loadings, nine emotional dimensions were identified and labeled accordingly. Factor 1 was labeled *Irritation*, reflecting emotions such as dissatisfaction and boredom, while Factor 2, characterized by distrust, despair, and anger, was labeled *Anger*. Factor 3 was labeled *Distress* as it encompassed emotions of disgust, sorrow, and shock, and Factor 4 was labeled *Anxiety*, capturing anxiety, fear, and surprise. Factor 5 was labeled *Tension*, indicating the contrast between expectancy and comfort. Factor 6, defined by exhaustion and realization, was labeled *Fatigue*. Factor 7 was labeled *Gratitude*, representing relief, respect, and gratitude, whereas Factor 8 was labeled *Care* and Factor 9 was labeled *Pride*. Because these emotional factors were extracted from the reactions of actual moviegoers, the factor loading values can be interpreted as indicating the extent to which each film possesses attributes capable of eliciting specific emotional responses.

The labeling of Factor 5 warrants particular theoretical justification, as it exhibits a bipolar dimension. Lehne and Koelsch (2015) define tension as "an emotional state characterized by a desire for change that is felt to eventually occur." In music, this relates to the expectation-driven experience of listeners. Building on this, Factor 5's bipolar structure—anticipation (+) versus contentment (−)—closely corresponds to foundational emotion models: Thayer's (1989) energetic arousal dimension, Russell's (1980) circumplex arousal axis, and Scherer's (2001) arousal-based categorization. In these frameworks, high- arousal emotions (e.g., anticipation) contrast with low-arousal emotions (e.g., serenity). From a psychological perspective, anticipatory tension is neither pure surprise nor joy—it is an expectancy-driven state associated with heightened physiological arousal (heart rate, cortisol levels). This aligns with Scherer's felt action tendency and Russell's arousal-based differentiation of affective states. Therefore, for movies—where temporal unfolding and suspense are narrative cornerstones—*Tension* is both a conceptually valid and psychologically grounded label for this bipolar factor.

3.3 Regression Analysis Results

OLS regressions examined the impact of the nine emotion factor scores on p, q, and m. The variance inflation factors (VIFs) of all independent variables included in the regression models were below 6, indicating that multicollinearity was not a concern.

In the regression analysis using the innovation coefficient (p) of the Bass model as the dependent variable (Table 3), Factor 1 (*Irritation*) exhibited a significantly positive effect, whereas Factors 2 (*Anger*), 5 (*Tension*), and 6 (*Fatigue*) showed significantly negative effects. This counter-intuitive finding—that *Irritation* drives innovation—can be interpreted through the Affect-as-Information lens. According to Schwarz and Clore (1983, 1996), negative emotions signal that the current state is problematic, activating systematic information processing. Audiences experiencing Irritation may interpret this emotion as a signal of novelty or unconventional storytelling, prompting early adoption. Critically, this effect was observed while controlling for structural factors (e.g., genre, 〈Table 3〉 OLS Regression Results (p)

Variables	Coefficient	Std. Error	t-value	p-value
Constant	.0884	.010	8.733	.000
Factor 1	.0176	.003	6.571	.000
Factor 2	-.0072	.003	-2.234	.027
Factor 3	-.0040	.002	-1.673	.096
Factor 4	.0023	.003	.712	.477
Factor 5	-.0086	.004	-2.162	.032
Factor 6	-.0082	.003	-2.950	.004
Factor 7	.0006	.003	.185	.854
Factor 8	-.0017	.003	-.505	.614
Factor 9	.0007	.003	.300	.765
Numposter	-.0016	.001	-2.889	.004
Numtrailer	.0011	.001	1.395	.165
Numscreen	-1.847e-05	5.47e-06	-3.379	.001
Summer	.0115	.006	1.883	.061
Winter	.0024	.008	.302	.763
Seollal	-.0295	.013	-2.342	.020
Chuseok	-.0007	.010	-.073	.942
N	209			
Adj. R^2	0.346			

Note: Nineteen dummy variables for film genres were included.

screens, seasonality), indicating that the specific emotional tone of a film—not merely its marketing scale—exerts a causal influence on adoption patterns.

Among the control variables, the number of poster images, number of screens, and the Seollal dummy variable were found to have significantly negative effects.

These results indicate that films with stronger attributes likely to elicit irritation-related emotions tend to attract a higher proportion of consumers who choose to watch the movie independently. In contrast, films with stronger attributes associated with anger, tension, or fatigue are associated with a lower proportion of self-initiated moviegoers. Moreover, contrary to expectations, an increase in the number of posters and screens was found to decrease the innovation coefficient, suggesting that, unlike traditional explanations, greater marketing communication efforts and wider distribution channel accessibility may actually reduce the proportion of consumers who decide to watch a film on their own initiative.

In the regression analysis using the imitation coefficient (q) of the Bass model as the dependent variable, Factor 2 (*Anger*) and Factor 5 (*Tension*) exhibited significantly positive effects. These results provide robust support for the Arousal-Driven Sharing and Social Sharing of Emotion theories. Both *Anger* and *Tension* represent high-arousal emotional states (Russell, 1980; Thayer, 1989). According to Berger (2011), high-arousal emotions—regardless of valence—create a

〈Table 4〉 OLS Regression Results (q)

Variables	Coefficient	Std. Error	t-value	p-value
Constant	.0193	.017	1.119	.265
Factor 1	-.0021	.003	-.674	.501
Factor 2	.0083	.003	2.996	.003
Factor 3	-.0010	.002	-.462	.645
Factor 4	-.0028	.003	-.997	.320
Factor 5	.0094	.003	2.750	.007
Factor 6	-.0032	.002	-1.320	.188
Factor 7	.0026	.003	.846	.399
Factor 8	.0028	.003	.956	.340
Factor 9	-.0030	.002	-1.375	.171
WOMvol	-1.532e-07	3.56e-07	-.430	.668
WOMval	.0011	.002	.530	.597
Summer	-.0013	.005	-.245	.807
Winter	.0005	.007	.077	.939
Seollal	.0820	.011	7.493	.000
Chuseok	.0256	.009	2.973	.003
N	209			
Adj. R^2	0.345			

Note: Nineteen dummy variables for film genres were included.

physiological readiness that motivates individuals to share the experience, whether through direct communication or online word-of-mouth. Rimé (2009) further posits that socially sharing emotional experiences helps individuals regulate emotions and seek social validation. In the context of movie diffusion, films that induce *Anger* or *Tension* generate conversational capital, accelerating word-of-mouth transmission. Crucially, these effects were observed while controlling for 19 genre dummies (e.g., Action, Thriller, Drama). This ensures that the imitation parameter (q) captures the marginal effect of a film's specific emotional execution—the particular *Tension* of this thriller—beyond the baseline word-of-mouth pattern of its genre. Thus, our results reveal that the emotional signature of a film has a causal, genre-independent effect on social diffusion. Among the control variables, both the Seollal and Chuseok dummy variables showed significantly positive effects.

These results suggest that films with stronger attributes likely to evoke *Anger* or *Tension* tend to attract a higher proportion of viewers who decide to watch the film under the influence of other consumers. Furthermore, consistent with expectations, films released during Seollal and Chuseok holidays exhibited higher imitation coefficients, indicating that during peak seasons, a larger number of low-involvement consumers—who are more susceptible to word-of-mouth—tend to watch movies.

In the regression analysis using the market potential (m) of the Bass model as the dependent variable, Factor 1 (*Irritation*) showed

〈Table 5〉 OLS Regression Results (m)

Variables	Coefficient	Std. Error	t-value	p-value
Constant	-2.551e+06	4.35e+05	-5.861	.000
Factor 1	-3.752e+05	1.15e+05	-3.265	.001
Factor 2	2.041e+05	1.38e+05	1.477	.141
Factor 3	1.452e+05	1.03e+05	1.409	.161
Factor 4	1.515e+05	1.37e+05	1.107	.270
Factor 5	3.486e+05	1.71e+05	2.039	.043
Factor 6	2.256e+05	1.2e+05	1.878	.062
Factor 7	-1.164e+05	1.49e+05	-.781	.436
Factor 8	-5.57e+04	1.47e+05	-.379	.706
Factor 9	-1.092e+05	1.08e+05	-1.014	.312
Numposter	3.618e+04	2.31e+04	1.567	.119
Numtrailer	-2.448e+04	3.36e+04	-.729	.467
Numscreen	3502.7224	235.174	14.894	.000
Summer	5.528e+04	2.62e+05	.211	.833
Winter	1.105e+05	3.38e+05	.327	.744
Seollal	3.034e+05	5.41e+05	.560	.576
Chuseok	-7.95e+04	4.28e+05	-.186	.853
N	209			
Adj. R^2	0.649			

Note: Nineteen dummy variables for film genres were included.

a negative effect, whereas Factor 5 (*Tension*) exhibited a significantly positive effect. Among the control variables, the number of screens was found to have a significantly positive effect.

These results indicate that films with stronger attributes likely to evoke irritation-related emotions tend to have lower market potential, whereas those that elicit tension-related emotions tend to exhibit higher market potential. In addition, consistent with previous findings on film-related determinants, the positive correlation between the number of screens and box office performance was reaffirmed.

3.4 Robustness Check

To assess the temporal stability of emotion expressions in online reviews, an additional analysis was conducted. For this check, review data were collected at three specific time points: the day of release (Day 0), seven days post-release (Day 7), and fourteen days post-release (Day 14). A cap of 1,000 reviews per day per movie was set,

〈Table 6〉 Chi-square Test Results

Comparison	Chi-square	p-value	df
Day0 vs. Day7	0.0086	1.000	37
Day7 vs. Day14	0.0295	1.000	37
Day0 vs. Day14	0.0518	1.000	37

and the analysis included 42 movies for which at least 300 reviews were available at each of these three time points.

Chi-square tests were performed to compare the distribution of the 39 emotion ratios across these time points (Day 0 vs. Day 7, Day 7 vs. Day 14, and Day 0 vs. Day 14). As shown in 〈Table 6〉, the results indicated no statistically significant differences in the overall emotion distributions between these early periods of a film's release. This suggests that the emotional tenor established by very early viewers tends to remain relatively consistent during the life cycle, supporting the use of early review emotions as indicators.

Ⅳ. Discussion and Conclusion

This study empirically examined the role of emotional attributes of films extracted from online reviews in shaping movie diffusion, as modeled by the Bass framework. By linking the Bass model parameters to emotional factors captured via a deep learning-based natural language processing approach (KcELECTRA), the research offers both theoretical insights and practical guidance. In particular, it highlights how emotional characteristics of films can serve as meaningful signals in predicting market outcomes in the evolving Korean film industry—especially under shifting consumer behaviors in the post-COVID era.

4.1 Theoretical Implications

The key contribution of this study lies in establishing an empirical link between the emotional attributes of films and their diffusion trajectories. Factor 1 (*Irritation*) and Factor 5 (*Tension*) showed opposite effects on the innovation coefficient and market potential—positive and negative for one, negative and positive for the other. This suggests that irritation-inducing attributes may attract innovators and boost early audience inflow but ultimately hinder box office performance, whereas tension-inducing attributes may have limited impact initially yet enhance long-term success. These contrasting patterns imply that innovators and the general audience evaluate emotional attributes using markedly different, even opposing, criteria.

Similarly, Factor 1 (*Anger*) and Factor 5 (*Tension*) negatively affected the innovation coefficient but positively influenced the imitation coefficient, indicating that films evoking clear anger targets or strong tension tend to be avoided by innovators but embraced by imitators. Collectively, these findings highlight that imitators play a decisive role in determining a film's ultimate commercial performance in the Korean market. Notably, Factor 5 (*Tension*) significantly affected all three Bass model parameters. Positively associated with expectancy and negatively with comfort, attraction, and happiness, this factor represents a cinematic attribute that evokes suspense through unpredictability. In essence, films that elicit tension and suspense are more likely to become the subject of word-of-mouth communication and achieve commercial success in the Korean film market.

From a production standpoint, filmmakers should consider how a film's emotional tone influences its diffusion trajectory. Since tension-related attributes promote long-term success while irritation-related ones mainly attract early audiences but deter broader viewership, creating balanced tension or controlled suspense may sustain audience engagement beyond the initial release. From a marketing perspective, the findings highlight that imitators play a decisive role in determining box office outcomes. Accordingly, post-release strategies should focus on word-of-mouth amplification and emphasize the emotional depth and suspenseful appeal of films, rather than mere novelty. Leveraging peak seasons and socially engaging content can further enhance diffusion among later audiences.

4.2 Limitations and Future Research

Several constraints warrant caution in interpreting the findings and point to opportunities for future research. First, the use of static summary statistics derived from early reviews—rather than dynamic, time-varying emotional signals—represents a methodological choice with important trade-offs. While this approach effectively captures the emotional tone perceived by the majority of viewers (especially due to the Law of Large Numbers effect with our large sample of 182,987 reviews), it does not account for temporal evolution in emotional expressions or shifts in audience composition over a film's lifecycle. Future research could employ state-space models or time-varying coefficient frameworks (e.g., Kalman filters) to explicitly model how emotional signatures evolve over time and how such dynamics influence diffusion parameters. Second, while our large sample size (183,000 reviews) ensures ecological validity by leveraging the Law of Large Numbers to capture genuine common sentiment rather than idiosyncratic noise, sample representativeness and data collection strategies also present limitations. The dataset consists of 209 Korean films released within a particular period, and reviews were exclusively gathered from a single cinema website (CGV). Emotional expression patterns may differ on other platforms (e.g., Naver or international streaming services). Third, although we controlled for 19 genre dummies to isolate the marginal effect of a film's specific emotional execution, perfect disentanglement of structural factors remains challenging. For instance, certain genres (e.g., Thriller) inherently predispose films toward higher *Tension* scores, making it difficult to fully separate genre-driven expectations from execution-driven emotional impact. Future studies could employ hierarchical or random-effects models to more rigorously decompose within-genre and between-genre variance. Fourth, the emotional classification relies on a KcELECTRA-based model fine-tuned with the KOTE dataset at a high prediction-score threshold (.8). Although this approach strengthens accuracy, similarities or ambiguities among certain emotions—such as boredom versus disappointment—remain challenging. Cultural and linguistic nuances specific to Korean online discourse further limit generalizability. Subsequent research could expand to multi-lingual contexts, refine labeling criteria, or consolidate highly similar emotions to reduce classification ambiguity.

References

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.

Berger, J. (2011). Arousal increases social transmission of information. *Psychological Science*, *22*(7), 891-893.

Berger, J., Moe, W. W., & Schweidel, D. A. (2023). What holds attention? Linguistic drivers of engagement. *Journal of Marketing*, *87*(5), 793-809.

Chevalier, J. A., & Mayzlin, D. (2006). The Effect of Word of Mouth on Sales: Online Book Reviews. *Journal of Marketing Research*, *43*(3), 345-354.

Choe, S., & Kang, J. (2017). CJ CGV's Entry into Vietnam Through a Cross-Border Acquisition. *Korea Business Review*, *21*(3), 113-145.

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.

Jang, Y., Choi, J., & Kim, H. (2022). KcBert-based movie review Corpus emotion analysis using emotion vocabulary dictionary. *J. KIISE*, *49*(8), 608-616.

Jeon, D., Lee, J., & Kim, C. (2022). *User Guide for KOTE: Korean Online Comments Emotions Dataset* (arXiv:2205.05300). arXiv.

Kim, D. J., Park, D. I., & Park, J. S. (2018). Study on the Change of Marketing Strategy through Data Mining Technique. *Korea Business Review*, *22*(2), 177-194.

Korean Film Council. (2024). *2023 Korean Film Industry Reports.*

Kwon, J. E., & Kim, S. (2021). CJ CGV's Strategies for Global Entry and Diffusion: Focusing on Special Screening Platforms. *Korea Business Review*, *25*(3), 1-33.

Lee, J. (2021). *KcELECTRA: Korean comments ELECTRA* [Computer software]. GitHub. https://github.com/Beomi/KcELECTRA

Lee, Y., Cha, K. C., & Kim, S.-H. (2017). Exploring the impact of distributor's decision making on movie diffusion: The case of Korean market. *Journal of Korean Marketing Association*, *32*(3), 25-44.

Lehne, M., & Koelsch, S. (2015). Toward a general psychological model of tension and suspense. *Frontiers in Psychology, 6*, 79.

Liu, Y. (2006). Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue. *Journal of Marketing, 70*(3), 74-89.

Ludwig, S., de Ruyter, K., Friedman, M., Brüggen, E. C., Wetzels, M., & Pfann, G. (2013). More than Words: The Influence of Affective Content and Linguistic Style Matches in Online Reviews on Conversion Rates. *Journal of Marketing, 77*(1), 87-103.

Mokryn, O., Bodoff, D., Bader, N., Albo, Y., & Lanir, J. (2020). Sharing Emotions: Determining Films' Evoked Emotional Experience from their Online Reviews. *Information Retrieval Journal, 23*(5), 475-501.

Rimé, B. (2009). Emotion elicits the social sharing of emotion: Theory and empirical review. *Emotion Review, 1*(1), 60-85.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*(6), 1161-1178.

Scherer, K. R., Schorr, A., & Johnstone, T. (Eds.). (2001). *Appraisal processes in emotion: Theory, methods, research.* Oxford University Press.

Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: informative and directive functions of affective states. *Journal of Personality and Social Psychology*, 45(3), 513-523.

Schwarz, N., & Clore, G. L. (1996). Feelings and phenomenal experiences. In A. Kruglanski & E. T. Higgins (Eds.), *Social psychology: Handbook of basic principles, 2*, 385-407.

Thayer, R. E. (1989). *The biopsychology of mood and arousal.* Oxford University Press.

Ullah, R., Alam, M. A., & Zeb, A. (2025). How Emotions in Online Reviews Affect Movie Sales: Evidence from Hollywood. *Journal of Retailing and Consumer Services*, 85, 104304.

Zhu, F., & Zhang, X. (Michael). (2010). Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics. *Journal of Marketing, 74*(2), 133-148.

영화의 감정적 속성이 확산 패턴에 미치는 영향 - Bass 모형을 활용한 탐색적 분석 -*

이지환** · 이유석*** · 김상훈****

요 약

본 연구는 온라인 리뷰 텍스트에서 추론한 영화의 감정적 속성이 국내 영화 시장의 확산 패턴에 미치는 영향을 탐색적으로 분석한다. 182,987건의 관객 리뷰에 딥러닝 기반 감정 분류 모델(KcELECTRA)을 적용하여 39개 감정 변수를 추출하고, 요인분석을 통해 9개 차원으로 축소했다. Bass 확산 모형으로 추정한 혁신계수(p), 모방계수(q), 잠재시장 규모(m)를 종속변수로 설정하고 회귀분석을 실시했다. 본 연구는 정보로서의 감정 이론(Affect-as-Information)과 각성 주도 공유 이론(Arousal-driven Sharing)에 기초하여, 특정 감정이 시장 행동으로 연결되는 심리적 메커니즘을 규명한다. 분석 결과, '긴장감'과 '짜증' 요인은 혁신계수와 잠재시장 규모에 상반된 효과를 나타냈으며, 특히, '긴장감'은 모방계수를 유의하게 증가시켜, 서스펜스나 예측 불가능성을 유발하는 영화가 구전 기반 관객을 확보하고 장기적으로 더 강한 성과를 달성할 가능성이 높음을 시사한다. 특히, 이러한 감정 효과는 장르 및 스크린수 등 구조적 요인을 통제한 후에도 여전히 유의한 것으로 나타났다. 본 연구의 결과는 감정 기반 텍스트 신호와 영화 확산 역학 간의 관계에 대한 새로운 실증적 근거를 제시하며, 포스트 팬데믹 시대의 변화하는 영화 산업에서 제작자와 마케터가 감정적 톤과 관객 참여 전략 간의 균형을 모색하는 데 기여할 것으로 기대한다.

주제어: 정보로서의 감정, 각성 주도 공유, Bass 확산 모형, 감정분석, 영화 산업, 자연어 처리, 온라인 리뷰, 감정의 사회적 공유

---

* 본 논문은 서울대학교 경영대학 경영연구소 연구비 지원에 의해 이루어졌습니다.
** 삼정KPMG, 제1저자
*** 동국대학교 경영대학 경영학과, 부교수
**** 서울대학교 경영전문대학원, 교수, 교신저자