Marketing Science

Frontiers: In-Consumption Social Listening with Moment-to-Moment Unstructured Data: The Case of Movie Appreciation and Live Comments

Qiang Zhang, Wenbo Wang*, Yuxin Chen

Please scroll down for article—it is on subsequent pages

Frontiers: In-Consumption Social Listening with Moment-to-Moment Unstructured Data: The Case of Movie Appreciation and Live Comments

Qiang Zhang,[a] Wenbo Wang,[b,*] Yuxin Chen[c]

[a] School of Management and Economics and Shenzhen Finance Institute, The Chinese University of Hong Kong, Shenzhen (CUHK-Shenzhen), 518172 Shenzhen, China; [b] Department of Marketing, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong; [c] Stern School of Business, New York University, New York, New York 10012
*Corresponding author

Contact: zhangqiang@cuhk.edu.cn, https://orcid.org/0000-0003-4613-3100 (QZ); wenbowang@ust.hk,
https://orcid.org/0000-0002-4288-0929 (WW); yc18@nyu.edu (YC)

Abstract. Consumption of entertainment products such as movies, video games, and sports events often lasts a nontrivial time period. During these experiences, consumers are likely to encounter temporal variations in the content of consumption, to which they may react in real time. Compared with existing in-consumption analysis (e.g., eye tracking and neural activity analysis), listening to in-consumption consumers' voices on social media has great potential. Our paper proposes a new approach for in-consumption social listening and demonstrates its value in the context of online movie watching wherein viewers can react to movie content with live comments. Specifically, we propose to listen to the live comments through a novel measure, moment-to-moment synchronicity (MTMS), to capture viewers' in-consumption engagement. MTMS refers to the synchronicity between temporal variations in the volume of live comments and those in movie content mined from unstructured video, audio, and text data (i.e., camera motion, shot length, sound loudness, pitch, and spoken lines). We demonstrate that MTMS significantly predicts viewers' postconsumption appreciation of movies and that it can be evaluated at a finer level to identify engaging content. Finally, we discuss the information value of MTMS with the presence of measures used in the previous literature and the value of integrating supply-side content information into in-consumption analysis.

1. Introduction

Entertainment products such as movies, video games, and sports events often last a nontrivial amount of time and display temporal variations in their content. Among both marketing practitioners and academic researchers, there is a growing interest in the collection and analysis of in-consumption data that document real-time consumer reactions to these content dynamics. For example, Disney recently invested over $1 billion in MagicBand, a digital wristband that can track customers' locations and consumption process (e.g., park rides and food purchases) in its theme parks.[1] Marketing research firms such as Nielsen have adopted a variety of technologies such as functional magnetic resonance imaging, electroencephalogram, and eye and face tracking to track and monitor in-consumption consumer reactions. These in-consumption

data or measures have broad and important implications for understanding the drivers of overall product evaluation (Ramanathan and McGill 2007), designing more effective advertisements (Baumgartner et al. 1997, Teixeira et al. 2010), and forecasting product demand (Barnett and Cerf 2017).

The recent development of online social media offers a new form of in-consumption data: live comments. Major live streaming websites have adopted a live chat function that enables consumers to converse in real time while streaming video content. Listening to the live social media content generated by online consumers offers great potential for in-consumption analysis. Live social media allows mass participation without researcher interference and provides a platform to examine various consumption occasions in realistic settings. Moreover, social media data can be

collected at a low cost, making them particularly useful to study long and dynamic consumption contexts.

This research, the first of its kind, proposes in-consumption social listening and demonstrates its value in the empirical context in which viewers can react to movie content in real time with live comments that are also visible to other viewers. Specifically, we propose to listen to these live posts through a novel measure, moment-to-moment synchronicity (MTMS). MTMS refers to the synchronicity between temporal variations in the content of consumption and moment-to-moment (MTM) consumer reactions to these variations. In entertainment products such as TV shows and movies, producers often deploy a series of temporal variations in the content (e.g., plot development, motion, music) to hold consumers' attention and keep them engaged. When the content successfully engages the consumers, it may shape consumers' live commenting behaviors, resulting in highly synchronized commenting patterns that match the content variations. When consumers get distracted or bored with the content, the synchronization may drop. Overall, a movie's ability to engage the viewers can be reflected by the extent to which the live commenting pattern synchronizes with the movie's content variations.

To demonstrate the proposed approach, we collect and merge a unique data set of movie videos and live comments that document at the moment-to-moment (MTM) level what viewers see, hear, and talk about during the entire movie streaming. First, we introduce how to empirically construct MTMS. Construction of MTMS is challenging because the data on movie content and live comments are both highly unstructured. For instance, movie videos are a combination of unstructured, high-frequency data streams (image sequence, audio, and text). In addition, movies are complex products in which the meanings of scenes are context dependent. We tackle these difficulties by following "film grammar," the grammatical principles generally followed by professional filmmakers (Arijon 1976, p. 4). For instance, in film grammar, a shot, the basic component of a movie, is a continuous camera action. Whereas shorter shots are often manipulated to highlight conflict and emphasize a particular movie moment, longer shots are used more for expositional purposes and in calmer scenes. Therefore, the length of shots during different scenes of a movie can reveal which parts are intended to be emphasized and highlighted. Following similar principles and applying a set of scalable video processing tools, we capture MTM variations in movie content by quantifying shot length, camera motion, sound loudness, sound pitch, and number of spoken lines. Finally, we take the MTM volume of live comments as the collective consumer reactions and then operationalize a movie's MTMS using the R^2 value

from a regression of comment volume onto movie content variations.

We demonstrate the information value of MTMS by showing that it can significantly improve the prediction of viewers' postconsumption appreciation of the movies. First, we find that the information value of MTMS persists under several alternative specifications and for two different measures of movie appreciation (i.e., consumer ratings and virtual currency awarded by viewers). Second, we show that MTMS contributes as a unique, important, and robust source of in-consumption information. The variation explained by MTMS is orthogonal to that explained by several measures adopted in the existing literature on in-consumption analyses and social listening. Third, we find a more pronounced MTMS effect among suspenseful movies, providing support for the role of MTMS in capturing viewer engagement (Green and Brock 2000). Finally, we show that MTMS can be evaluated at different periods during an experience and for different types of reactions, leading to a refined view of viewer engagement.

Conceptually, our research views the consumption of an experience as a continuous communication process from the supply-side content producers to the demand-side consumers (Lang 2000). In this vein, MTMS can be seen as a metric for the effectiveness of information delivery during this process and can have broader implications for the analyses of other in-consumption context (or data) in which similar communication processes take place.

Managerially, MTMS is a valuable metric that helps to predict viewers' appreciation of content and to track viewer engagement during different periods of an experience (e.g., different parts or moments during a movie or TV show). For content suppliers such as movie or TV show producers, test screening is often employed to collect audience feedback before actual release of the content.[2] By analyzing live feedback from a test audience, producers may use MTMS to measure audience engagement during a movie or TV show. Such information is useful not only for postproduction decisions such as content editing but also for promotional activities such as selecting content for movie previews and trailers. For video platforms or TV networks, many shows are often released sequentially.[3] By analyzing live feedback from early episodes, companies can identify what type or part of the content is attractive and edit future episodes accordingly. MTMS can also be used to improve the timing of midroll ads. Midroll ads are likely to interrupt viewer experiences given that they are often inserted with fixed time intervals regardless of video content. The MTMS metric may help video platforms adjust timing of midroll ads to minimize the negative consequence of ad interruptions. Meanwhile, it is also

possible to use MTMS to improve the effectiveness of midroll ads, given that engagement in media content is a key factor that influences consumers' perceptions about in-show advertising (Wang and Calder 2006). Finally, by helping to predict the ratings and other performance measures of video content, MTMS can be used in content selection and recommendation by online video platforms.

Our research contributes to the emerging literature on analyzing MTM in-consumption data (e.g., neural measures, eye tracking, and facial expression tracking). Existing research has extracted summary statistics to predict product performances and advertising effectiveness (e.g., Baumgartner et al. 1997, Barnett and Cerf 2017) or has performed functional data analysis to extract granular information within MTM data (e.g., Hui et al. 2014). Unlike existing in-consumption analyses, which have focused on consumer-generated MTM data alone, our research contributes by proposing a novel measure, MTMS, which combines supply-side MTM content dynamics with demand-side MTM data.

MTMS also adds to the literature on intersubject correlation (ISC) that has received substantial attention in neuroscience and marketing (Hasson et al. 2004, Barnett and Cerf 2017). Specifically, in their seminal work, Hasson et al. (2004) find that individual subjects display highly synchronized brain activity when watching the same movie clip separately, but lower ISC is observed during an unedited video. Importantly, the synchronization appears in brain areas that govern high-level processing (e.g., selective attention and emotional response), suggesting that well-crafted content can control viewers' interpretations (Hasson et al. 2004, Barnett and Cerf 2017). The analysis of ISC also has been extended to other type of MTM data—for example, in the concentration of subjects' eye fixation positions during commercials (Teixeira et al. 2010) and in the coherence of self-report ratings (Ramanathan and McGill 2007). Relative to ISC, high MTMS requires not only high ISC among individual consumers but also high consumer-side synchronization with the product content. In this sense, MTMS can be more informative than ISC, which is supported by our empirical analysis.

Finally, our research also contributes to the broad theme of social listening (e.g., Godes and Mayzlin 2004, Seiler et al. 2017) by examining a new and fast-growing type of live social media that occurs at the MTM level during product consumption.

2. Live Commenting and MTMS
On live streaming sites, the live chat function enables consumers to chat with each other online while watching the broadcasted content. It has gained great popularity.[4] Our research focuses on a popular variant

of this function: the live commenting function. Figure 1 illustrates its format. The live commenting function enables viewers to post comments through an interactive box below the video player while streaming a video. These live comments are recorded in sync with video content; that is, each of them is attached to a specific time in the video. When another viewer watches the video later, these comments will appear at the top of the screen at the same moment in the video when the commenters originally post them. In this way, the viewer can see and react to previous viewers' comments.[5]

The live commenting function has been widely adopted by major online video platforms in China in nearly all video categories and has attracted mass viewer participation.[6] The live commenting function documents real-time viewer reactions to video content, offering a new form of in-consumption data. In Online Appendix A, we provide a case study to illustrate key features of live comments.

In this research, we conceptualize media consumption (e.g., watching movies or TV shows) as a continuous communication process in which producers deploy a series of temporal content variations (e.g., motion, music, and plot development) to engage viewers (Lang 2000). The successful delivery of a narrative experience often manifests as the active participation of the viewers (Bezdek et al. 2013). For example, Alfred Hitchcock described a movie scene in which two characters are sitting at a table under which a bomb is ticking, but they are unaware of the situation (Truffaut 1984). Viewers might long to warn the characters, "You shouldn't be talking about such trivial matters. There is a bomb beneath you, and it is about to explode!" Such active participation often indicates the viewers' strong narrative engagement such that they are acting as if they are inside the scenario described in the narrative (Green and Brock 2000). However, in the field setting of live commenting, live comments can be generated in response to not only the video content as a form of active participation but also environmental factors or random cues (Berger 2014), especially when the content fails to interest consumers or hold their attention.

Accounting for these key features, we propose a measure, MTMS, to listen to the live comments. Specifically, MTMS refers to the synchronicity between volume of live comments and movie content variations at the MTM level.[7] Whereas the volume of live comments indicates viewers' participatory behaviors, the movie content dynamics represent the variations intentionally created by content producers during a movie. Therefore, MTMS describes the degree to which the viewers' active participation is successfully controlled by the filmmakers and thus reflects the ability of the content to engage viewers.

Figure 1. (Color online) Example of Video Player with the Live Commenting Function
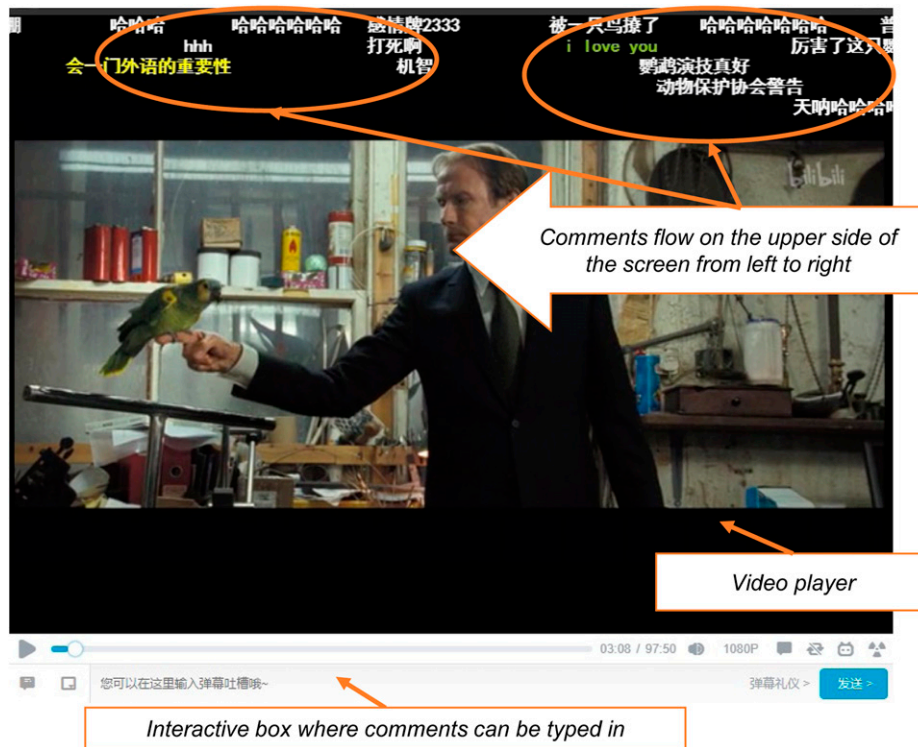
In the following sections, we explain how to extract movie content variations and construct MTMS. Because in-consumption engagement is closely related to postconsumption viewer appreciation, we further show that MTMS can be used to improve the prediction of viewer appreciation.

3. Data Overview, Processing, and Description
3.1. Data Overview
We collected and constructed a rich data set of 507 movies consisting of movie video, live comment data, and movie characteristics from a major video platform in China (for data collection details, see Online Appendix B). The live comment data contain the calendar time when the comment was posted, the time during the movie's run time when the comment was posted, the viewer's identification, and the content of each comment. The movie video file contains the stream data of image frames and audio. We collect the subtitle file, which reveals the timing and content of all the spoken lines. Note that images, audio, subtitles, and live comments are stream data at the MTM level, capturing what the viewers see, hear, and discuss during a movie.

In addition to the stream data, we collected the movies' static characteristics, including (1) consumer ratings from Douban Movie, a leading online movie review and critic site in China; (2) the movie's release

date, length, genre, and main actors; and (3) the popularity of the movie on the video platform, including total number of views, live comments, and coins (a type of virtual currency awarded to a movie by viewers). To capture the impact of movie star presence, we code an actor as a *star* if any of the movies he or she stars in has a rating above 7.5 on the Internet Movie Database.[8] Regarding movie genres, we focus on the most common genres and code them as dummy variables.[9] These movie characteristics are reported in Table 1.

3.2. Extracting Movie Content Variations
There are two challenges in quantifying movie content from raw videos. First, movie videos are a combination of unstructured, large, and high-frequency data streams. For instance, the total run time of all the movies in our sample is more than a month, with images available at over 24 frames per second and audio stream available at over 40,000 Hz. The second challenge relates to the complexity of movie products. A typical movie often covers many different contexts, characters, objects, and backgrounds. It is difficult to determine which context is more engaging than the other. Moreover, visually similar scenes in a movie can convey different meanings depending on the context. For example, in a scene in which two people are talking to each other in front of a table, they could be having a friendly conversation or they could be arguing. Although the director intends to deliver the tension in the latter case,

Table 1. Description of Movie Characteristics ($N = 507$)

	Mean	SD	Minimum	Maximum
Movie rating	6.52	1.18	2.7	9.2
Coin count	213	421	2	5,922
View count	177,397	262,551	11,919	5,035,928
Live comment count	6,043	5,711	2,005	59,396
Run time (min)	100	14.04	71	170
Release year	2010	4.14	2000	2015
Star actors	1.45	1.41	0	5

Note. Genres: Action (17%), comedy (33%), drama (41%), horror (35%), thriller (28%), romance (19%).

one may not be able to distinguish between these two visually similar scenarios just from recognizing characters or background of a movie scene. Therefore, a scalable and effective way is required for computers to read the movie content.

To tackle the above-mentioned difficulties, we identify and quantify major content dimensions following the literature on filmmaking techniques (film grammar) and computer science. Film grammar refers to a set of cinematic conventions or guidelines for film producers on how to employ editing techniques, camera movements, sound, and character interactions to direct viewers' attention and deliver intended meaning (Arijon 1976, Adams et al. 2002, Bordwell and Thompson 2013). Guided by the film grammar, we extract major movie content variations that are commonly created during film production. First, in film grammar, a *shot*, the basic element of a movie, is a continuous camera action. Whereas shorter shots are often manipulated to highlight conflict and emphasize a particular movie moment, longer-duration shots typically reflect calmer and more expositional narration in a movie (Cutting et al. 2010, Cutting 2016). Therefore, we first quantify the shot length of each shot during a movie. The second visual dimension we quantify is *motion*, which refers to object and camera movement. In film grammar, higher motion is often used to hold viewers' attention and highlight conflict during a movie moment (Adams et al. 2002, Cutting 2016). In addition, perceptual properties of a film's sound can shape the experience of the film (Bordwell and Thompson 2013). Thus, we extract the two most important properties of film sound: *sound loudness* and *pitch*. Finally, conversations between characters play an important role in film narration (Cutting 2016). We count the number of spoken lines in the subtitles as a measure of informational variation in a movie's narration. In summary, we quantify the temporal variations in movie content by shot length, camera motion, sound loudness, sound pitch, and number of spoken lines.

Although there are other content dimensions that relate to filmmakers' intention in creating movie

content dynamics, we focus on the five content dimensions mentioned earlier because they are major components in film grammar that can be automatically extracted with relatively high accuracy and low computing cost.[10] It takes roughly one-fourth of the movie's actual play time to extract the movie content variables with a desktop personal computer (PC).[11] Details of the video data processing are provided in Online Appendix B.

3.3. Description of Movie Stream Data
We divide each movie into T ($T = 100$) equal time intervals and aggregate the stream data into these intervals (see also Hui et al. 2014).[12] Figure 2 describes the distributions of live comment volume and movie content variables.[13] All these variables show substantial variations along the movie time. Comparing different genres, we find that horror movies have the highest average comment volume, whereas romance movies have the lowest. Interestingly, action, horror, and thriller movies use high motion levels, short camera shots, and low-pitched, high-energy sounds, whereas comedy, romance, and dramatic movies have more spoken lines. This is consistent with movies of different genres: to induce a feeling of thrill and excitement, horror, thriller, and action movies place more emphasis on audiovisual effects and less on conversations between characters relative to dramatic, comedy, and romance movies.

4. Empirical Analysis of MTMS
4.1. Operationalization of MTMS
MTMS refers to the level of synchronicity between a movie's live comment volume over time and the temporal variations in movie content. Empirically, we construct MTMS of a movie i with R^2 of regressions

$$\ln(Vol_{it} + 1) = \beta_{i,0} + X_{it} \cdot \beta_{i,1} + e_{it}, \qquad (1)$$

where Vol_{it} refers to the volume of live comments,[14] and X_{it} is a vector of movie content variables at the tth time interval of movie i after aggregation. Specifically, we include average values of shot length, number of spoken lines, sound loudness, sound pitch, and motion level.[15]

The regression R^2 captures the percentage of variation in volume of live comments that is explained by the variation in movie content. It captures how closely comment volume follows the movie content and has the following desirable features. First, R^2 does not restrict the direction of variables' influence. For example, during intense movie scenes, viewers may be highly engaged and not posting comments, creating a negative correlation between some of the movie content variables and the comment volume. In addition, R^2 can capture this type of negative correlation as well.

Figure 2. Description of Movie Stream Data Along Movie Time by Genre
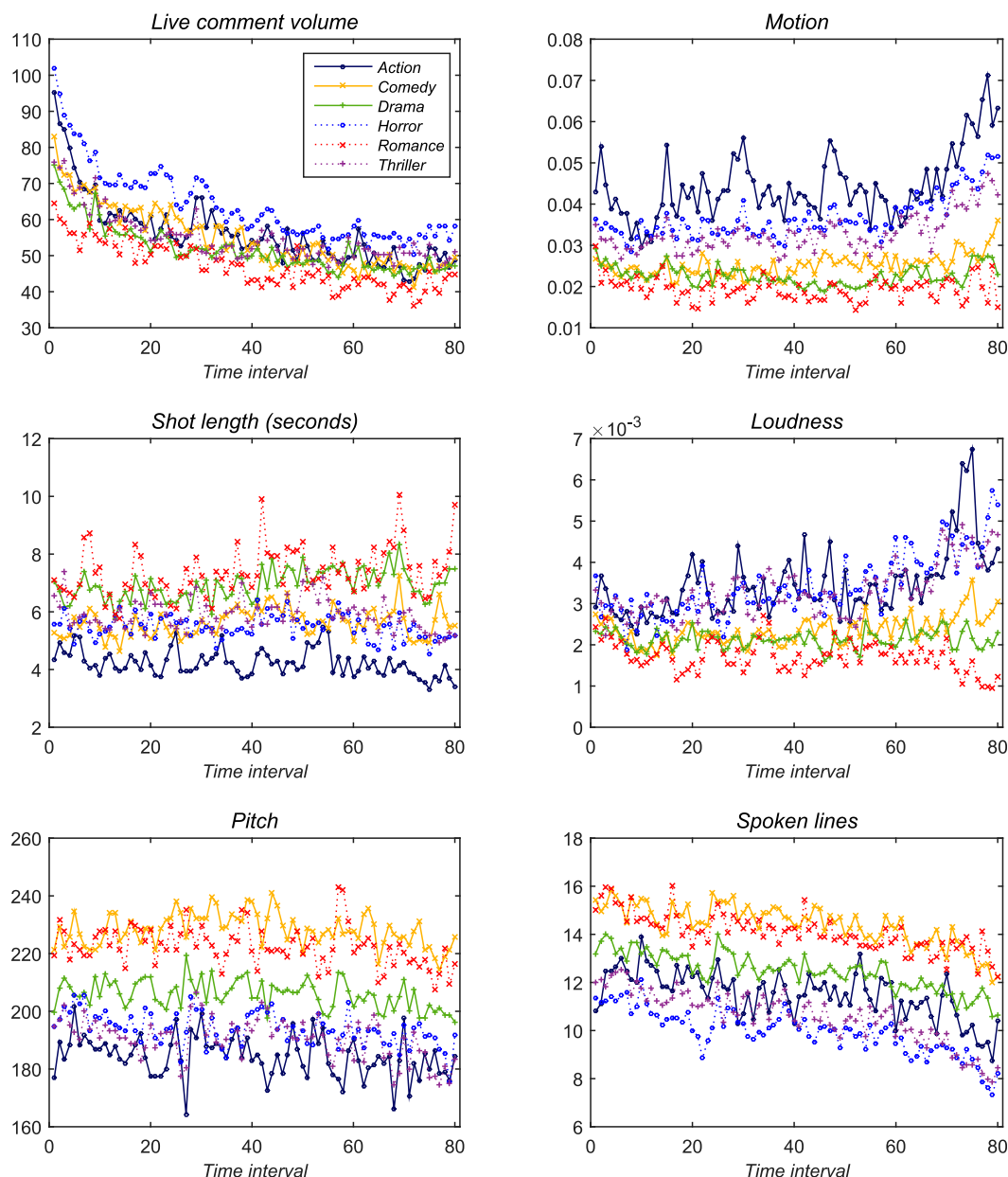
Notes. We remove data points during the first 5 and last 15 time intervals that are opening and ending credits. Therefore, for each movie, the number of observations is 80. This figure displays the average value of live comment volume and movie content variables among movies from each genre in these time intervals.

Second, R^2 can handle multiple content dimensions and can be measured for different periods of a movie. Finally, R^2 builds on regression models that can flexibly capture consumers' commenting behaviors.[16]

Our use of R^2 is inspired by an influential measure in the finance literature, stock price synchronicity, that describes how closely company prices synchronize with the market-level stock price. It is often constructed with the R^2 from the regression of company-specific prices on market-level price index (Morck et al. 2000). Price synchronicity is widely used

to measure how much market-level information (versus idiosyncratic company information) is impounded in company prices in a sector or the overall stock market. Here we use R^2 to describe the extent to which the content variations created by the producers are reflected in the consumers' MTM reactions or, more broadly, the effectiveness of information transmission from content suppliers to content consumers during the consumption process.

When constructing MTMS, we extract the earliest 2,000 live comments (according to calendar time) each

movie received to address the following concerns.[17] First, when overall comment volume changes, the distribution of live comment volume may be different, and this may also affect the construction of MTMS. Second, because live comments are displayed on the screen while the movie is playing, the viewers who posted the earliest 2,000 live comments experienced relatively little distraction.[18] Finally, focusing on the earliest comments is also managerially plausible because video platforms can use MTMS to investigate viewer engagement at an early stage after a video is released.

The value of R^2 obtained from the within-movie regressions, or MTMS, ranges from 0.011 to 0.552, with standard deviation (SD) close to 0.1, showing considerable variation across movies. Online Appendix B.4 provides a summary of these within-movie regressions.

4.2. MTMS and Viewer Appreciation

We measure viewers' appreciation of a movie in two ways. On the video platform, users can earn virtual coins by staying active (e.g., logging in and sharing videos). After watching a movie, they can "upvote" the movie with one coin to show their appreciation. We use the logarithm of aggregate number of coins awarded to a movie ln(*Coin*) as the first measure of viewer appreciation. In addition, we obtain a public movie rating *Rating* from Douban Movie.

To examine our central research question—whether MTMS can predict movie appreciation—we estimate the following cross-sectional regression:

$$MovieApprec_i = \gamma_0 + \gamma_1 MTMS_i + \boldsymbol{CtrlVar_i} \cdot \boldsymbol{\gamma_2} + w_i, \tag{2}$$

where $MovieApprec_i$ refers to *Rating* or ln(*Coin*), $MTMS_i$ is the synchronicity measure, $\boldsymbol{CtrlVar_i}$ contains a vector of control variables, and w_i is the error term. In terms of movie characteristics, we include a movie's release year, run time, and all the genre dummies as control variables. We also control for the means and SDs of the movie content variables.[19] In addition, to show the informational value of MTMS, we need to control for other variables that are potentially related to movie quality and may be correlated with both viewers' engagement and post-consumption evaluation. For this purpose, we include the number of star actors and volume of live comments as control variables. However, because using overall comment volume may disguise the effect of movie content variables, we create and include a variable, excessive volume, that captures the volume of live comments for a movie beyond (either higher or lower than) the average level predicted by the movie content variables (see Online Appendix C).[20] Finally, because ln(*Coin*) is largely affected by movie popularity at the platform, we include the logarithm of the number of views of a movie at the platform in the regression of ln(*Coin*).[21]

The first two columns of Table 2 display the regression results, which show that MTMS plays a significant role in explaining both measures of viewer appreciation ($p < 0.01$). All else being equal, a 0.01 increase in MTMS is associated with 1.3% increase of virtual coins and a 0.01-point increase in the 10-point scale of movie rating. In terms of predictive power, including MTMS contributes to a 0.6% and 0.8% increase in R^2 in the regression of *Rating* and ln(*Coin*), respectively. Table 2 also shows that movies with a large amount of excessive volume receive high appreciation. In addition, we find that the effect of MTMS remains similar when we exclude excessive volume, suggesting that the data variations in viewer appreciation explained by MTMS are orthogonal to the excessive volume. We further present a series of robustness tests with different regression specifications used to measure MTMS in Online Appendix D.

A concern with these results is that viewers' enjoyment during a movie can also be driven by their social interactions through the live comments (e.g., Ramanathan and McGill 2007). However, this is unlikely to be the main driver because the existing live comments and the interactions they enable are unlikely to be independent of the movie content. In most cases, only after the movie content induces certain comments do the viewers start to interact with these comments. Moreover, one of the dependent variables, *Movie Rating*, is generated by viewers from a separate platform, and these viewers are unlikely to have interacted with the live comments while watching the movie. Thus, if viewer interaction were the main driver, we would not see the significant positive relation between MTMS and movie rating.

As for control variables, we find that the means and SDs of movie content variables do not significantly explain *Rating* or ln(*Coin*), except that movies with longer shots have significantly higher ratings. No significant result is shown in the regression of ln(*Coin*). Next, older movies have a higher *Rating* than newer movies, but no significant year effect is found in ln(*Coin*).[22] In addition, longer movies and movies with more star actors receive significantly higher viewer appreciation. Finally, horror and action movies receive significantly lower appreciation, whereas dramas receive higher appreciation.

In addition, we expect the effect of MTMS to be stronger among movies in which suspense and tension are key elements (see Green and Brock 2000).[23] The ability of these movies to engage viewers can be particularly important. We run the same analysis with movies from the horror and thriller genres and find a more pronounced effect of MTMS (see columns (3) and (4) of Table 2).[24] Increasing MTMS by

Table 2. Effects of MTMS on Viewer Appreciation

	(1)	(2)	(3)	(4)
Sample	All	All	Horror and thriller	Horror and thriller
	Rating	ln*(Coin)*	*Rating*	ln*(Coin)*
MTMS	1.021***	1.303***	1.983***	1.734***
	(0.374)	(0.423)	(0.497)	(0.567)
Excessive volume	0.00419***	0.0112***	0.00355***	0.00998***
	(0.000789)	(0.00108)	(0.00107)	(0.00133)
Motion mean	0.650	−3.161	1.721	−1.434
	(3.807)	(4.628)	(4.381)	(6.401)
Shot length mean	0.0438**	−0.00629	0.0593*	−0.0131
	(0.0206)	(0.0277)	(0.0343)	(0.0369)
Lines mean	0.00363	0.00493	0.0218	0.00639
	(0.0133)	(0.0148)	(0.0192)	(0.0243)
Motion SD	−1.193	4.487	2.248	4.799
	(4.688)	(5.362)	(5.979)	(8.077)
Shot length SD	−0.0311	−0.0197	−0.0319	−0.0124
	(0.0210)	(0.0249)	(0.0319)	(0.0340)
Lines SD	0.00152	−0.0229	−0.0737	−0.0624
	(0.0323)	(0.0398)	(0.0488)	(0.0630)
Year	−0.0782***	0.00318	−0.0783***	0.0126
	(0.00818)	(0.00994)	(0.0117)	(0.0144)
Run time	0.0106***	0.00761**	0.0259***	0.0183***
	(0.00307)	(0.00359)	(0.00540)	(0.00616)
Star actors	0.165***	0.118***	0.150***	0.0874
	(0.0276)	(0.0316)	(0.0445)	(0.0538)
Action	−0.517***	−0.0175	−0.451***	0.0714
	(0.126)	(0.124)	(0.158)	(0.155)
Comedy	0.0339	0.0909	0.328**	0.376*
	(0.0941)	(0.105)	(0.157)	(0.193)
Drama	0.409***	0.247**	−0.00694	−0.0617
	(0.0970)	(0.102)	(0.120)	(0.137)
Horror	−0.740***	−0.885***	−0.748***	−1.071***
	(0.116)	(0.113)	(0.147)	(0.172)
Romance	0.126	0.0605	0.0478	0.183
	(0.0925)	(0.117)	(0.343)	(0.676)
Thriller	−0.110	−0.0347	0.00612	−0.0556
	(0.0952)	(0.101)	(0.132)	(0.139)
ln*(View)*		0.219***		0.256***
		(0.0729)		(0.0945)
Constant	162.3***	−5.250	160.7***	−25.31
	(16.46)	(19.88)	(23.59)	(28.91)
Observations	507	507	254	254
R^2	0.576	0.532	0.567	0.540
Adjusted R^2	0.561	0.515	0.535	0.505

Note. Robust standard errors in parentheses.
  *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

0.01 increases coins by 1.7% and results in approximately a 0.02-point increase in movie rating, all else being equal. The contribution to R^2 by MTMS amounts to 2.8% in movie rating and 1.7% in number of coins.[25] This subsample analysis provides further support for the role of MTMS in capturing viewer engagement.

4.3. Diagnostic Analysis with MTMS

Beyond analysis at the movie level, MTMS can be built at different periods of a movie and can provide diagnostic information regarding which part of a movie contributes to the main source of overall viewer appreciation. In analyzing film narratives, Thompson (1999) divides film structure into four acts: setup,

complication, development, and climax. Each act lasts for approximately one-quarter of a movie's run time. Following this notion, we construct MTMS for each movie quartile and include the quartile-specific MTMS in place of the overall MTMS in the movie-level regressions.[26] The results are reported in Table 3. Among all four quartiles, the last is the most important in predicting *Rating* and ln(*Coin*). This result is in line with the end effect that ending experience has a disproportionate influence on overall evaluation (e.g., Hui et al. 2014). Our analysis complements these studies and offers additional evidence in a field setting.

MTMS can also be constructed with different types of consumer reactions in the live comments. We show that only MTMS constructed with volume of content-related comments is significant in explaining both *Rating* and ln(*Coin*), suggesting that the MTMS effect is mainly driven by the comments that explicitly mention movie content rather than by communication or sentiment-related comments (see Online Appendix E). This finding further suggests that social interaction is not the main driver of the result.

4.4. The Robust Information Value of MTMS

After establishing the predictive value of MTMS, we further show (in Online Appendix F) that MTMS contains unique and robust information that cannot be captured by other measures frequently adopted in existing in-consumption analyses and the social listening literature. These measures include (1) functional principle components, which capture evolution patterns of live comment volumes; (2) peak–end comment volumes; (3) volume of engagement-related comments; (4) volume of sentiment-, content-, and communication-related comments; (5) volume of comments on specific sentiments; and (6) (empirical) topic assignment from unsupervised or supervised topic models. These analyses indicate that MTMS contributes as a unique and important source of in-consumption information.[27]

We also conduct an analysis controlling for ISC, which is a measure that has received substantial attention recently (Hasson et al. 2004, Ramanathan and McGill 2007, Teixeira et al. 2010, Barnett and Cerf 2017). Whereas the basic idea behind the predictive value of ISC is that well-crafted content can create synchronized MTM reactions among individual consumers, MTMS goes beyond this because it combines the dynamics of supply-side product content with the demand-side consumer reactions. Specifically, in the case of live commenting, high synchronization between viewers (i.e., ISC) results in an amplified comment volume peak or trough, but a movie will have a similar ISC measure no matter when the comment volume reaches peaks or troughs during the movie. In contrast, high MTMS requires comment volume peaks

Table 3. Effects of MTMS on Viewer Appreciation Measured for Different Movie Quartiles

	(1)	(2)
	Rating	ln(*Coin*)
MTMS first quartile	−0.130	0.453
	(0.408)	(0.438)
MTMS second quartile	0.264	0.534
	(0.433)	(0.504)
MTMS third quartile	−0.452	−0.340
	(0.403)	(0.482)
MTMS fourth quartile	0.761*	1.016**
	(0.404)	(0.487)
Observations	507	507
R^2	0.571	0.531
Adjusted R^2	0.553	0.510

Notes. Robust standard errors in parentheses. Control variables are identical to the main regressions.
　*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

and troughs to appear at specific moments defined by movie content (to be synchronized with movie content variations). We follow Barnett and Cerf (2017) to construct ISC with our live commenting data. After controlling ISC in the movie regression, we still find a robust MTMS effect, and the ISC measure is not significant. The contribution of R^2 by MTMS is also larger than ISC, which indicates that MTMS is more effective in predicting viewer appreciation than ISC. Details of these empirical analyses can be found in Online Appendix F.

5. Conclusion and Future Research

This research is the first to propose in-consumption social listening and demonstrate its value in the empirical context where viewers can post live comments during online movie watching. We develop a novel measure, MTMS, to capture viewer engagement during content consumption. Our empirical analysis demonstrates MTMS as an important, unique, and robust source of in-consumption information and provides further insights by evaluating MTMS during different periods of an experience and for different types of consumer reactions. The proposed measure not only has implications for content producers and online platforms but also provides a general approach to analyzing MTM consumer reaction data.

As the first paper on in-consumption social listening, our research has only scratched the surface of this exciting area, and there are many opportunities for future research. First, although we focus on the movie category, the methods proposed here can be generally applied to other consumption scenarios and other types of in-consumption data. In the subsample analysis, our MTMS metric shows more significant effect in movie categories such as horror and thriller

than in other genres. This can be seen as a falsification test of the intuition behind MTMS because most of our movie content variables are about audiovisual effects that are likely to play more important roles in horror and thriller movies than in nonsuspenseful movies (e.g., dramatic and romance movies), which rely more on the dramatic delivery of emotional scenes. To extract such content variations from these non-suspenseful movies, more sophisticated computing tools may be needed. We acknowledge this possibility and call for future research on genre-specific movie content analysis to generate movie content variables used for measuring MTMS.

Second, our empirical analysis centers on the operationalization and validation of the MTMS concept and demonstrates that MTMS can predict ratings and awards given by viewers after consumption. Additional research is needed to examine whether and how MTMS drives product demand and product-related word of mouth. Future research can also investigate the downstream application of MTMS in the decisions of online video platforms such as optimizing the effectiveness of midroll ads.

Finally, we show that in-consumption measure MTMS is an important predictor of video appreciation by viewers. Future research can use MTMS in combination with machine learning approaches to improve the selection and recommendation of video content on video platforms.

Acknowledgments

Endnotes

[1] See https://www.nytimes.com/2014/04/02/business/billion-dollar -bracelet-is-key-to-magical-kingdom.html (accessed January 31, 2019).

[2] During test screenings, producers often screen the movies to selected groups of people for comments and suggestions. Directors may change the presentation of content or even the storyline based on the feedback. For an example of the changes made on *Pacific Rim Uprising* (2018) after test screenings, see https://screenrant.com/pacific-rim -uprising-changed-test-screenings/amp/ (accessed August 20, 2019). More examples can be found at https://en.wikipedia.org/wiki/ Test_screening (accessed August 20, 2019).

[3] As shown in our empirical analysis presented later in this paper, MTMS can be constructed with only the earliest 2,000 live comments and available soon after a video's release on the video platform.

[4] Through the live chat function, 14.2 billion chat messages were sent on Twitch TV in 2016, a 55% increase from 2015; see https:// www.twitch.tv/year/2016/ and https://www.twitch.tv/year/2015/ (accessed March 7, 2018).

[5] The interaction among viewers takes place through anonymous comments; viewers' identities are not shown. To illustrate this commenting function, we provide a YouTube video about the format of live comments and how to post them (available at https:// youtu.be/W7Vw6gPWmQQ).

[6] For example, as of February 2017, the first episode of a popular Chinese TV show on Tencent Video, *Ode to Joy II* (2017), has received over 2.2 million live comments.

[7] Viewers may post negative comments that may not reflect their active participation, and they may delay their commenting during engaging movie periods. We systematically address these issues in the robustness analysis. For details, see Online Appendix D.

[8] In this practice, we use a larger movie list from the MovieLens data set ($N = 18,703$). For the MovieLens data set, see https://grouplens.org and Harper and Konstan (2016). We obtained similar results when we change the rating cutoff point to seven or eight.

[9] In the data, there is no primary genre for a movie, and most of the movies are associated with two to four genres, which are listed in alphabetical order (e.g., "action comedy" or "comedy drama").

[10] Some previous research in film study relied on human coders to identify film editing and making techniques by inspecting movie videos (see Cutting et al. 2010, Cutting 2016). We follow the computer science literature (e.g., Apostolidis and Mezaris 2014) to automatically extract filmmaking techniques from video files, which is more suitable in our case given the large data size and our intention of making the proposed approach more generalizable and implementable. To provide additional support for the choice of movie content dimensions, we employ the movie content variables to predict movie genres with logistic regressions. The hit rates range from 75.7% to 88.6% across movie genres. The analyses are available upon request.

[11] We used a desktop computer with Intel Core i7-4790 central processing unit with a clock speed of 3.6 GHz and a random-access memory size of 16 GB to process the video data.

[12] We present data aggregation details in Online Appendix B. For data aggregation, we also tried a different aggregation level ($T = 200$) for the movies, and the results still hold.

[13] We remove data points during opening and ending credits (i.e., the first 5 and last 15 time units) that are unrelated to the movie content. For each movie, the number of observations is $N = T - 5 - 15 = 80$. We tried a specification with removing the final 10% of each movie instead of 15% and obtained similar results.

[14] We take the logarithm of comment volume because its distribution is skewed.

[15] Sound variables (sound loudness and pitch) are normalized by means and SDs because people can adjust sound features while watching movies.

[16] We run robustness checks with respect to the regression models that include lagged content variables, autoregressive components, and nonlinear movie content variables. They all show a robust MTMS effect. See Online Appendix D for details.

[17] All the movies in the sample have more than 2,000 live comments.

[18] We tried an alternative specification, using the earliest 1,500 live comments each movie received. The result is similar to that reported here.

[19] Because sound loudness and pitch do not have variations across movies after normalization, we do not include them in the movie-level regression.

[20] When using the logarithm of overall live comment volumes instead of the excessive volume in the movie regression, we find very similar results.

[21] Ideally, we need to control for how many times the movie has been watched by viewers who can give a coin. On the video platform, only registered viewers can give coins to movies. Unfortunately, the data reveal only the total number of views by registered and

nonregistered viewers. Therefore, we use the total number of views in the regression. We also tried an alternative specification with probability of giving a coin to the video among all the viewers as dependent variable. The results remain similar.

[22] The different results in the regression of *Rating* versus ln(*Coin*) are not surprising. Douban Movie, where we collected the movie ratings, is often considered a movie review site, and its users may account for the artistic value and historical significance of a movie when giving their ratings, whereas this is not likely the case for the video platform. See https://alltechasia.com/will-clamping-down-on-online-critics-improve-the-quality-of-chinese-films/ (accessed December 25, 2017).

[23] Green and Brock (2000) show that narrative transportation, a strong form of engagement in narrative content, can impact audiences' evaluations and their beliefs. In the scale used to measure transportation, Green and Brock use feelings of suspense as one of the major components, in addition to emotional involvement in the story, cognitive attention to the story, lack of awareness of surroundings, and mental imagery. In our context, given that suspenseful movies are to create suspenseful feelings, viewers' feelings of suspense can be appreciated more in this genre relative to others. Therefore, if MTMS can capture the viewers' engagement or their feelings of suspense, MTMS is expected to have a stronger predictive power in movie appreciation.

[24] We do not use a regression model with the interaction term between MTMS and movie genres. Because of the small sample sizes at the genre level, we are unable to run reliable regressions for specific genres or consider a multilevel model to capture the different effects of MTMS as well as the control variables across genres.

[25] The role of MTMS is highlighted by movies that do not have high appeal (e.g., movies with few star actors or low budget) but deliver an immersive experience. For example, *The Call* (2013) is a thriller with a high rating of 7.3 out of 10 and MTMS as high as 0.46. Whereas the model without MTMS predicts a rating of 6.22 for this movie, the model including MTMS improves the prediction to 6.70. Another well-known psychological thriller, *The Butterfly Effect* (2004), has a rating of 8.7 and is predicted to have a rating of only 8.02 without MTMS; however, the prediction rating improves to 8.39 when we include MTMS.

[26] We divide each movie into four quartiles. For each quartile, we construct the MTMS by dividing the movie within that quartile into 80 time intervals of equal length. We chose 80 to be consistent with the main analysis.

[27] It can be costlier to compute MTMS compared with other metrics mainly because computation of MTMS involves processing video data. The computational cost is still affordable, however. For example, it takes roughly 25% of the actual playtime to process a movie video with a desktop PC. The processing time can be even shorter considering the higher computing power owned by video platforms or movie studios. On a Google cloud server, we find that processing a 95-minute high-definition movie only takes about 15 minutes. In addition, one can speed up the processing by paralleling the processing of multiple videos and using videos with relatively lower definition.

References

Adams B, Dorai C, Venkatesh S (2002) Toward automatic extraction of expressive elements from motion pictures: Tempo. *IEEE Trans. Multimedia* 4(4):472–481.

Apostolidis E, Mezaris V (2014). Fast shot segmentation combining global and local visual descriptors. *Proc. Internat. Conf. Acoustics Speech Signal Processing (ICASSP)* (IEEE Computer Society, Washington, DC), 6583–6587.

Arijon D (1976) *Grammar of the Film Language* (Silman-James Press, Hollywood, CA).

Barnett SB, Cerf M (2017) A ticket for your thoughts: Method for predicting content recall and sales using neural similarity of moviegoers. *J. Consumer Res.* 44(1):160–181.

Baumgartner H, Sujan M, Padgett D (1997) Patterns of affective reactions to advertisements: The integration of moment-to-moment responses into overall judgments. *J. Marketing Res.* 34(2):219–232.

Berger J (2014) Word of mouth and interpersonal communication: A review and directions for future research. *J. Consumer Psych.* 24(4):586–607.

Bezdek MA, Foy JE, Gerrig RJ (2013) "Run for it!": Viewers' participatory responses to film narratives. *Psych. Aesthetics Creativity Arts* 7(4):409–416.

Bordwell D, Thompson K (2013). *Film Art: An Introduction*, 6th ed. (McGraw-Hill, New York).

Cutting JE (2016) Narrative theory and the dynamics of popular movies. *Psych. Bull. Rev.* 23(6):1713–1743.

Cutting JE, DeLong JE, Nothelfer CE (2010) Attention and the evolution of Hollywood film. *Psych. Sci.* 21(3):432–439.

Godes D, Mayzlin D (2004) Using online conversations to study word-of-mouth communication. *Marketing Sci.* 23(4):545–560.

Green MC, Brock TC (2000) The role of transportation in the persuasiveness of public narratives. *J. Personality Soc. Psych.* 79(5):701–721.

Harper FM, Konstan JA (2016) The MovieLens datasets: History and context. *ACM Trans. Interactive Intelligence Systems* 5(4):19:1–19:19.

Hasson U, Nir Y, Levy I, Fuhrmann G, Malach R (2004) Intersubject synchronization of cortical activity during natural vision. *Science* 303(5664):1634–1640.

Hui SK, Meyvis T, Assael H (2014) Analyzing moment-to-moment data using a Bayesian functional linear model: Application to TV show pilot testing. *Marketing Sci.* 33(2):222–240.

Lang A (2000) The limited capacity model of mediated message processing. *J. Comm.* 50(1):46–70.

Morck R, Yeung B, Yu W (2000) The information content of stock markets: Why do emerging markets have synchronous stock price movements? *J. Finance Econom.* 58(1):215–260.

Ramanathan S, McGill AL (2007) Consuming with others: Social influences on moment-to-moment and retrospective evaluations of an experience. *J. Consumer Res.* 34(4):506–524.

Seiler S, Yao S, Wang W (2017) Does online word-of-mouth increase demand? (and how?) evidence from a natural experiment. *Marketing Sci.* 36(6):838–861.

Teixeira TS, Wedel M, Pieters R (2010) Moment-to-moment optimal branding in TV commercials: Preventing avoidance by pulsing. *Marketing Sci.* 29(5):783–804.

Thompson K (1999) *Storytelling in the New Hollywood: Understanding Classical Narrative Technique* (Harvard University Press, Cambridge, MA).

Truffaut F (1984) *Hitchcock* (Simon & Schuster, New York).

Wang J, Calder BJ (2006) Media transportation and advertising. *J. Consumer Res.* 33(2):151–162.