Received 26 February 2014
Revised 5 October 2014
13 January 2015
29 March 2015
Accepted 16 June 2015

Are online communities on par with experts in the evaluation of new movies? Evidence from the Fandango community

Pradeep Kumar Ponnamma Divakaran
Marketing Department, ESC Rennes School of Business, Rennes, France, and
Sladjana Nørskov
Department of Business Administration, Aarhus University, Aarhus, Denmark

Abstract

Purpose – The purpose of this paper is to investigate two questions. First, are movie-based online community evaluations (CE) on par with film expert evaluations of new movies? Second, which group makes more reliable and accurate predictions of movie box office revenues: film reviewers or an online community?

Design/methodology/approach – Data were collected from a movie-based online community Fandango for a 16-month period and included all movies released during this time (373 movies). The authors compared film reviewers' evaluations with the online CE during the first eight weeks of the movie's release.

Findings – The study finds that community members evaluate movies differently than film reviewers. The results also reveal that CE have more predictive power than film reviewers' evaluations, especially during the opening week of a movie.

Research limitations/implications – The investigated online community is based in the USA, hence the findings are limited to this geographic context.

Practical implications – The main implication is that film studios and movie-goers can rely more on CE than film reviewers' evaluation for decision making. Online CE can help film studios in negotiating with distributors, theatre owners for the number of screens. Also, community reviews rather than film reviewers' reviews are looked upon by future movie-goers for movie choice decisions.

Originality/value – The study makes an original contribution to the motion picture performance research as well as to the growing research on online consumer communities by demonstrating the predictive potential of online communities with regards to evaluations of new movies.

Keywords Evaluation, Decision making, User participation, Virtual community, Web 2.0

Paper type Research paper

1. Introduction

Although many different factors influence a movie's market performance, critics and audiences seem to be vital to this relationship. Indeed, Eliashberg *et al.* (2006) note that "the fickle reactions of audiences and critics make it difficult to accurately predict revenues and profits; each movie is unique" (p. 642). In addition, the emergence of movie-based online communities potentially changes the role which audiences play in a

The authors would like to thank Lars Esbjerg and Lars Frederiksen for their useful comments and suggestions to the early version of this paper. Similarly thanks to researchers at SCANCOR (Stanford University), and also the user and open innovation community, who provided useful inputs in shaping the ideas. Also, the authors thank Helle Alsted Søndergaard for her useful insights.

Information Technology & People
Vol. 29 No. 1, 2016
pp. 120-145

0959-3845
DOI 10.1108/ITP-02-2014-0042

movie's market performance. Audiences have benefited from the emergence of such communities, which has increased the amount of readily available movie-related information. As many industries shift their attention to consumer online communities (Kozinets, 1999; Pitta and Fowler, 2005), the role of online communities (movie-based) in the motion picture industry has never been explored before even though the number of movie-based online communities is growing rapidly. Thus in this study we investigate the potential role of a movie-based online community in the evaluation of new movies.

The role of film critics has already received substantial attention. Prior research has shown a link between market performance of movies and the role of film critics; a high degree of agreement has been established that film critics are significant as predictors of box office revenue (Eliashberg and Shugan, 1997; Gemser *et al.*, 2007; Reinstein and Snyder, 2005). Studies have further investigated the relationship between critic and novice agreement on movie quality (Plucker *et al.*, 2009), and the relationship between critic and layperson agreement on movie quality (Boor, 1992; Holbrook, 1999). Here, results show a correlation (although low) among critics and non-critics movie ratings (Holbrook, 1999; Plucker *et al.*, 2009). Our study aims to add to this pool of research by investigating the similarity (are they on par?) and prediction ability (which one is a better predictor of box office revenues?) of the movie reviews done by experts and those done by online communities.

Online communities are best described as a social relationship aggregation facilitated by internet technology in which members communicate and build personal relationships (Hagel III and Armstrong, 1997; Kozinets, 2002; Rheingold, 1993). Prior studies have shown that online communities foster trust and knowledge sharing among its members (Hara and Hew, 2007; Nambisan and Baron, 2010; Nolan *et al.*, 2007; Phang *et al.*, 2009), and provide a platform for social support (Leimeister *et al.*, 2008).

An online community is composed of interactive and non-interactive members (Burnett, 2000). Members exhibiting non-interactive behaviour observe other community members passively (e.g. reading other members' comments) and then participate in collective processes by utilising aggregation tools/mechanisms such as rating, ranking or voting. For example, in communities such as Flixster, Dell IdeaStorm and My Starbucks Idea, non-interactive members participate in voting or rating but do not post any comments. According to Carroll and Rosson (1996), the lurkers-to-posters ratio runs as high as 100:1 in most online communities. Posters are members who post comments and interact with other members. Lurkers are those who merely browse the community website and read messages posted in the discussion forums (Madupu and Cooley, 2010).

Prior studies related to problem solving in online communities have focused on either the interactive members (e.g. studies on open source software communities) or individuals (e.g. studies on lead user innovation), whereas the combined efforts of interactive and non-interactive members were largely ignored. We address this gap by focusing on the combined effect of the individual efforts of non-interactive members and the collective efforts of interactive members, which can take place simultaneously in an online community. For example, in a movie-based community called Fandango (www.fandango.com), the movie rating that is displayed is the overall average of the interactive members' ratings and the non-interactive members' ratings. In this study, the unit of analysis is not just the interacting members or the non-interacting members of an online community but both of them combined at the community level[1].

Using the motion picture industry and a movie-based online community, Fandango, as the empirical setting for this study, we investigate the following questions: are Fandango

community member evaluations on par with expert evaluations? And which group makes more reliable and accurate predictions of movie box office revenues: film reviewers[2] or Fandango community member evaluations? We do this by comparing expert (i.e. journalistic film reviewers') evaluations of movies with evaluations provided by the selected online community Fandango. We exclude those professional experts (e.g. the Academy Awards selection committee members) who judge the aesthetic or artistic merit or the value of movies, since the primary intention of our study is to evaluate movie enjoyment value which is crucial for their box office success rather than the artistic value of movies.

The remainder of this paper is organised as follows: Section 2 provides the theoretical background for our study by presenting the relevant research on the prediction of box office revenues and by identifying the conditions necessary for online communities to perform accurate and reliable evaluations of new movies. Section 3 discusses the identification of a suitable empirical setting for our study and explains the research methodology used in this study. Section 4 provides findings derived from the statistical analyses and their interpretation. Finally, Section 5 presents conclusions, implications, limitations and future research directions.

2. Theoretical background

2.1 Evaluation of new movies and box office performance

A film reviewer is a person who is usually employed by a media outlet to screen newly released movies and to provide potential audiences with comments and opinions regarding the movie (Cones, 1992 in Eliashberg and Shugan, 1997, p. 70). Scholars underline the difference between reviewing and criticism (Holbrook, 1999; Shrum, 1991). Professional critics "assess movies' artistic success from a relatively detached and long-term perspective that focuses on accepted standards for excellence" (Holbrook, 1999, p. 148). Conversely, journalistic reviewers work with "the day-to-day task of recommending certain movies to readership and viewership" members whose taste the reviewers may be eager to satisfy (Holbrook, 1999, p. 148). Thus, professional critics base their reviews on "aesthetic criteria associated with field-specific training or education", whereas journalistic reviewers provide recommendations based on their evaluations and predictions of the tastes and tendencies of their readers (Holbrook, 1999, p. 148). In this study, and unlike some previous studies (e.g. Holbrook, 1999), we focus only on the expert evaluation provided by journalistic reviewers[3] who evaluate a movie's quality in terms of its commercial value and popularity. It makes sense to study journalistic reviewers because they are, by definition, supposed to be knowledgeable about moviegoers' tastes and should thus be able to predict box office revenues. Indeed, one would expect a high level of agreement in movie evaluations of journalistic reviewers and moviegoers.

Film critics usually get a sneak preview or special screenings of the movie before its theatre release to enable them to publish their review ahead of the movie premiere. Some readers or subscribers look upon film critics as opinion leaders with special knowledge and expertise related to movies (Boatwright *et al.*, 2007; Eliashberg and Shugan, 1997; Turnball and Meenaghan, 1980; Venkatraman, 1989). However, studies have also questioned the role of film critics and their ability to influence moviegoers (Adams, 2006; Basuroy *et al.*, 2003; D'Astous and Touil, 1999; Hsu, 2006; Johnson, 2003). Influencers are defined as those critics who "influence their audience and thus the box office" through voicing their opinions, while predictors act as "a leading indicator of their respective audience with no significant influence on actual box office revenues" (Duan *et al.*, 2008,

p. 1010). Eliashberg and Shugan (1997) show that film critics are leading indicators of box office performance but their influencing ability is limited. One reason is that there are several other subjective factors (e.g. favourite actor, genre, storyline, awards, etc.), which outperform the influence of film critics' reviews. Another possible reason is that film critics understand their readers very well and their likes and dislikes, which is why film critics are good predictors of performance of new movies without influencing them (Basuroy *et al.*, 2003; Boatwright *et al.*, 2007; D'Astous and Touil, 1999).

Based on data from Dutch newspapers, Gemser *et al.* (2007) have investigated the impact of film critics' reviews on art-house movies and mainstream movies, and found that film reviews have an influence effect on art house movies and a predictor effect on mainstream movies (i.e. in predicting box office revenues). This predictor effect is similar to findings of other studies (such as, Basuroy *et al.*, 2003; Ravid *et al.*, 2006; Reinstein and Snyder, 2005), even though they are based on different market data such as the US and European data. This means that in spite of some cultural differences between US and European markets, the investigated relationship (i.e. film critics' predictive effect) is similar in the European and American markets.

Duan *et al.* (2008) inquired into the influence effect of online user reviews on movie sales and found no significant effect. Their study showed that, similar to film critics, online user ratings are predictors of movie performance but they do not influence movie performance. The authors conclude that "consumers are fully capable of inferring the true quality of a movie from online reviews without being influenced by the ratings of the reviews *per se*" (Duan *et al.*, 2008, p. 1008). This current knowledge, however, still leaves open questions about the similarity (are they on par?) and prediction ability (which one is a better predictor of box office revenues?) of the two groups of reviewers (i.e. film reviewers or movie-goers); this is what our paper aims to examine.

2.2 Online communities – evaluation of movies

Collective evaluation and selection are part of the community decision-making process (e.g. in deciding whether to watch a new movie) such that the output of any such decision-making process is a collective choice (Denning and Hayes-Roth, 2006; Oral *et al.*, 1991; Satnam, 2008; Sunstein, 2006; Surowiecki, 2004). As member participation behaviour in online communities can be interactive or non-interactive (Burnett, 2000; Madupu and Cooley, 2010), evaluation of new movies occurs in two manners. First, an evaluation may be based on the aggregation of all of the individual opinions of non-interactive members. This process is called "filtering" (Abony, 1983; Goldenberg *et al.*, 1992; Jabeur and Martel, 2007; Kim *et al.*, 2010; Luo, 2002). Second, an evaluation can be based on agreement or consensus (Iandoli *et al.*, 2007; Oral *et al.*, 1991) of the interactive members. Figure 1 summarises the above-mentioned steps that are involved in performing evaluation and selection tasks in the case of new movies.

Based on these prior works, we define new movie evaluation in online communities as the aggregation scores of all individual evaluations (regardless of any discussion with other community members) obtained via aggregation toolkits such as ratings (e.g. a five-point scale), voting (e.g. yes or no) and user reviews, to arrive at the community's majority preference.

2.3 Identifying the conditions required for online communities to reliably perform evaluation tasks

Understanding a community's decision-making process is a precursor to identifying the conditions that online communities should meet for the reliable performance of new

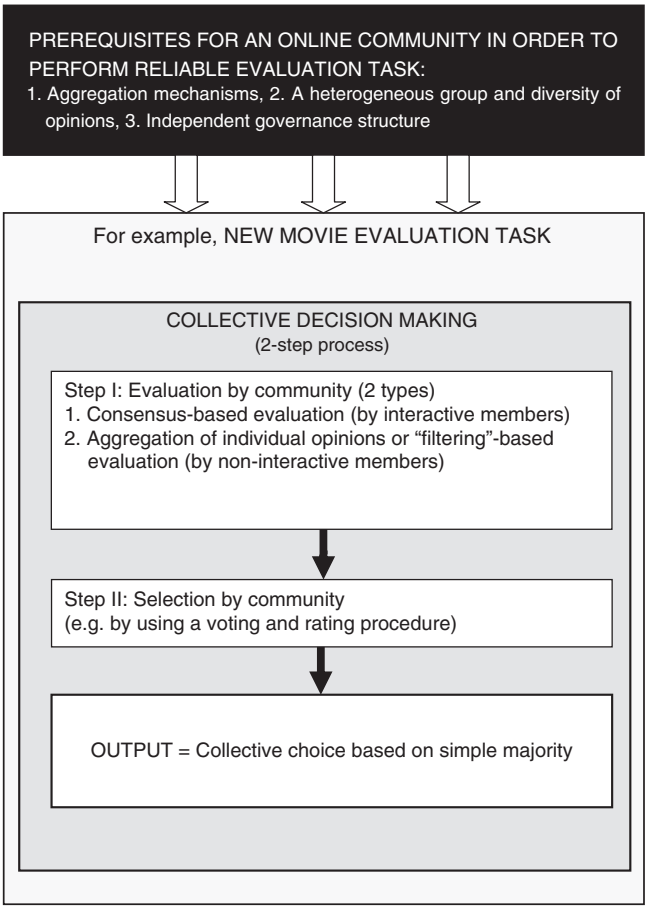

Figure 1.
Evaluation of new movies by an online community

movie evaluation tasks. Moreover, identifying these conditions also enabled us to identify a suitable online community for our study.

2.3.1 Aggregation mechanisms. Different online communities use different forms of aggregation mechanisms and tools. The community designers and founders usually decide which are appropriate, depending on the purpose or task to be performed by the community (Franke and Piller, 2004; Lin and Lee, 2006; Nambisan, 2002). For example, an aggregation tool, such as a five-point scale on the community website, provides a collective choice output in terms of "average ratings", which is an aggregate of all of the individual member ratings. Similarly, a voting mechanism provides an aggregate of the many members who voted for or against a particular discussion topic. Such tools enable each individual member to express or to share his opinion, expectation, satisfaction and so on. The aggregation tools thus capture the heterogeneity of user tastes and preferences thereby facilitating the determination of collective choice (Franke and von Hippel, 2003; Luo, 2002).

2.3.2 Heterogeneous groups and diversities of opinions. Extant research underscores the relevance of heterogeneity (in terms of personal experience,

localised knowledge and independent opinions) in reliable collective task performance (Franke and von Hippel, 2003; Kobayashi, 2010; Luo, 2002). In the context of community decision making, three additional interrelated criteria are thus emphasised (Surowiecki, 2004): first, the diversity of opinions based on each individual's private information and personal experience. If a group lacks diversity of opinion (e.g. a homogenous group), it may lead to polarisation, which is a group's tendency to radicalise their position on an issue (Sunstein, 2006). Second, each individual's independent opinion which are not influenced by the opinions of those around them, and third, the individual's ability to draw on local knowledge (referred to as decentralisation). If not, a common knowledge effect occurs whereby people tend to stick to the information and knowledge they already have and do not explore the diverse private stores of information that each member may have (Sunstein, 2006). When these criteria are fulfilled, any group of people will be able to make intelligent (collective) decisions better than any single individual would be able to do on his own (Sunstein, 2006). According to Surowiecki (2004), "if you ask a large group of diverse, independent people to make a prediction or estimate a probability, and then average those estimates, the errors each of them makes in coming up with an answer will cancel themselves out" (p. 10). Sunstein (2006) pointed out another cause called 'error amplification' which can fail a group in making accurate, truthful and the reliable decisions. Error amplification is determined by two major forces called informational pressure and social pressure. Informational pressure causes imitations when people tend to just believe things if a large number of other individuals purportedly do so, the rationale being "how can all these people be wrong?" Social pressure occurs due to peer influence such as fear of being ridiculed or marginalised by the group, low status, compliance to hierarchy and opportunity behaviour. Due to the conformism produced by these two forces, groups are unable to explore a large number of possible solutions and tend to converge too rapidly on a decision which is too obvious or a less-problematic one (Sunstein, 2006). Thus people may tend to imitate and the errors will spread and the probability that the group comes to bad decisions increases significantly leading to cascade effects causing error amplification and premature convergence (Sunstein, 2006).

Heterogeneity is identified as one of the conditions necessary for collective intelligence. However, community member heterogeneity is also important if the community is to be a good representation of the mass market. Heterogeneity is important in this manner because, according to the MPAA (2010)[4], moviegoer (market) demographics reveal that frequent (once a month or more) moviegoers represented 11 per cent (or 35 million people) of the population in 2010, and this relatively small group of frequent moviegoers is the locomotive of the industry, as they are responsible for more than 50 per cent of ticket sales. Similarly, among MPAA-rated movies, PG-13 films comprised 12 of the top 25 films in release during 2010. PG films continued to be popular; nine of the top 25 films were rated PG. The MPAA's report also reveals that 21 per cent of the total population of moviegoers was in the 25-39 age groups, 15 per cent were in the 40-49 age groups, 14 per cent were in the 2-11 age groups, 14 per cent were in the 50-59 age groups, 8 per cent were in the 12-17 age group, and 18 per cent were in the 60+ age group for the year 2010. The top three age groups of frequent moviegoers were the 25-39 age groups (22 per cent), the 18-24 age groups (21 per cent) and the 12-17 age groups (18 per cent). Comparing these market demographics with the community demographics revealed a high

degree of similarity, suggesting that this community might be a good representation of the mass market.

2.3.3 Independent governance structure. A virtual community is an independent, i.e. a self-managed and self-organised, system with a minimum of top-down management and centralisation (Iandoli *et al.*, 2007). For this type of organisation to function properly, the community must attract a critical mass of motivated, diverse users (attention governance) to reduce the risk of informational pressure, social pressure and common knowledge to prevent premature convergence and to enable the sufficient exploration of the search space. The community must determine its rules and organisational structures in terms of the process and roles that enable attention governance and effective participation (Iandoli *et al.*, 2007, pp. 6-7). When online communities meet these conditions, they are able to make reliable decisions without any internal or external pressure or influence, and thus live up to their independent governance structure.

Having identified the minimum conditions for the performance of collective decision-making tasks within online communities, we hypothesise that under the (aforementioned) three conditions, online communities can perform evaluation tasks that are accurate and reliable.

3. Research methodology

3.1 The selected online community

An online community can be either country specific or global and hence failure to take the different levels of data into account will lead to wrong conclusions (e.g. if the independent variables such as community evaluation (CE) and expert evaluation are US-based and if market performance data (box office revenues) is for worldwide collections, then it will clearly lead to wrong conclusions (e.g. due to cultural factors) since data are not at the same level of analysis). All the variables used in our study are very specific to the American (USA) movie industry, and hence avoids any bias due to country-level cultural factors. The Fandango community is US-based and all the movies focused in this study are released in US theatres. The box office revenues are for the US market only. Only US-based film critics and their reviews are included in this study. The movie online community Fandango (www.fandago.com) was used in this study as this community was found to meet all of the conditions (identified in Section 2.3) that would enable it to accurately and reliably perform the task of collective evaluation.

The members of the Fandango community are heterogeneous (a condition identified in Section 2.3.2) in terms of gender, age group, movie-watching frequency and individual preferences, tastes and opinions on different movie genres and MPAA ratings, all of which were demonstrated by the high standard deviation in the community members' movie ratings (SD was 0.74 on a five-point scale, mean was 3.44, min. was 1 and max. was 5 for 367 movies (six missing values) released during a 16-month period) and by the variety in their reviews and comments (55 per cent of all user comments were positive, 27 per cent were negative and 32 per cent were mixed for the 368 movies (five missing values) released during the 16-month study period). Moreover, in the Fandango community, different movie genres attract varying levels of member participation. The community demographics include both genders and all age groups (children, teens, adults and senior citizens), a fact that is clearly reflected when examining movies according to genre. Moreover, the community members represent

both frequent and infrequent moviegoers, and this can be noticed in the individual profile pages of the community members. Therefore, there is a diversity of opinions at the community level because of the diversity of preferences and tastes that the individual community members exhibit.

Such diversity in members' opinions and preferences leads to in-depth discussions before arriving at a final evaluation and hence prevents any premature conclusion. Participation in the community is voluntary and independent of external influence, i.e. independent governance structure, of any individuals or institutions (a condition identified in Section 2.3.3). In addition, the community provides aggregation mechanisms (for displaying collective choice), such as a rating scale, a voting mechanism (adoption intention), a comments section and information on the total number of members who participate for each movie (a condition identified in Section 2.3.1). Moreover, Fandango attracts a large number of members (on average, 714 members per movie) who write comments and rate each movie (SD was 1,690 members, min. was 0 and max. was 20,370), thus fulfilling a condition identified in Section 2.3.3.

3.2 Data collection, key variables and measures

The data were collected for a 16-month period, from 1 September 2009 until 31 December 2010, and included all movies released during this 16-month period (373 movies). For each movie released during this period, community-level data were collected on a weekly basis for up to eight weeks after the movie release. Sources for the different data used in this study are shown in Table I. The following sub-sections provide a detailed explanation of the operationalisation of the variables.

3.2.1 CE. Three different types of CE data were collected such as CE_average ratings, CE_percentage of positive evaluation and CE_percentage of negative evaluations (Table I). CE_average ratings data for each movie are the average of all of the five-star ratings given by community members after watching a movie. CE_average ratings data are collected on a weekly basis for up to eight weeks after a movie is released and also its lifetime average rating data are collected. In addition, the community (Fandango) provides information about the number of members who gave a five-star, a four-star, a three-star, a two-star and a one-star rating for each movie. Using this information the ratings and reviews were categorised into three types such as positive ($\geqslant$ 4-star ratings), negative ($\leqslant$ 2-star ratings) or mixed (between two and four-star ratings). From this data, the percentage of positive (and negative) CE is calculated (explained in Table I under the column "description").

3.2.2 Expert evaluation (film reviewers' evaluation or FRE). Metacritic.com was used as the source for film reviewers' evaluation data based on each movie's metascore. Each movie receives one or several reviews and Metacritic.com gathers all the review information for each movie from various sources. Subsequently, each film reviewer's input is given a value ranging from 1 to 100, and then the average of all the values for each movie is calculated. Thus, a metascore is a value given to each movie; it is the weighted average of the most respected film reviewers' online and print reviews. A higher (lower) metascore value suggests that the quality of movie is very good (bad). For example, the metascore value for the movie "The Social Network" (released on 1 October 2010) at 95 per cent was quite high, suggesting that it has good entertainment value and it is a good movie to watch. Another movie "Transylmania" (released on 4 December 2009) received a low metascore of just 8 per cent, suggesting

Variables	Description	Numerical scaletype	Name of sources	Links
1. Box office revenues (natural log)	Box office sales of each movie were collected on a weekly basis for 8 weeks, and the gross revenues during its lifetime were also collected. This is a dependent variable used in the analysis presented in Tables VI and VII	Ratio scale	1. The numbers[a] 2. ERC box office 3. Variety 4. IMDb databases	www.the-numbers.com www.ercboxoffice.com www.variety.com www.imdb.com
2. Film reviewers' evaluation (FRE)				
2.1. FRE average ratings (5-point scale)	The average of all the ratings given by different film reviewers for each movie. This variable is used in the analysis presented in Tables II and IV	Interval scale	1. Metacritic	www.metacrtic.com
2.2. FRE_Percentage positive reviews	=(Number of positive reviews/total number of positive, negative and mixed reviews)×100. This variable is used in the analysis presented in Table VI	Ratio scale	2. Metacritic	www.metacrtic.com
2.3. FRE Percentage negative reviews	=(Number of negative reviews/total number of positive, negative and mixed reviews)×100. This variable is used in the analysis presented in Table VII	Ratio scale	3. Metacritic	www.metacrtic.com
3. Community evaluation (CE)				
3.1. CE_Average ratings (5-point scale)	The average of all the ratings given by community members for each movie. This variable is used in the analysis presented in Tables II-V	Interval scale	1. Fandango	www.fandango.com
3.2. CE_Percentage positive reviews	=(Number of positive reviews/total number of positive, negative and mixed reviews)×100. This variable is used in the analysis presented in Table VI	Ratio scale	2. Fandango	www.fandango.com
3.3. CE_Percentage negative reviews	=(Number of negative reviews/total number of positive, negative and mixed reviews)×100. This variable is used in the analysis presented in Table VI	Ratio scale	3. Fandango	www.fandango.com
4. CE total number of reviews (natural log)	The total number of reviews posted by community members (which includes positive, negative and neutral reviews) for each movie. This variable is used in the analysis presented in Tables VI and VII	Ratio scale	1. Fandango	www.fandango.com
5. FRE_Total number of reviews (natural log)	The total number of reviews written by film reviewers (which includes positive, negative and neutral reviews) for each movie. This variable is used in the analysis presented in Tables VI and VII	Ratio scale	1. Metacritic	www.metacritic.com
6. Number of screens (natural log)	This data is collected on a weekly basis for each movie and also for the lifetime of each movie. This variable is used in the analysis presented in Tables VI and VII	Ratio scale	1. The numbers	www.the-numbers.com
7. Production budget (natural log)	This variable is used in the analysis presented in Tables VI and VII	Ratio scale	1. The numbers[a] 2. IMDb databases	www.the-numbers.com www.imdb.com

Table I.
Variable description and sources

Note: [a]Primary data source

that its entertainment value is low and it is not a movie worth watching. For data analysis purposes we transformed the percentage scale into a five-point scale by dividing the percentage values by 20. This transformation of film reviewers' evaluation into a five-point scale is carried out to ensure that the level of measurement of film reviewers' evaluation is the same as CE which also uses a five-point scale.

The average number of film reviewers who wrote reviews for each movie was 24 (from a total sample of 347 movies), with a minimum of one review (for the movie "Career", released on 4 September 2009) and a maximum of 42 reviews (for the movie "Harry Potter and the Deathly Hallows: Part 1", released on 19 November 2010). Metacritic.com also provides details such as the number of positive, negative and mixed reviews received for each movie.

3.3.3 Market performance. The market performance (dependent variable) was measured in terms of box office revenue (following prior studies such as e.g. Basuroy *et al.*, 2003; Eliashberg *et al.*, 2007; Gemser *et al.*, 2007; Liu, 2006; Sawhney and Eliashberg, 1996). The first eight weeks after a movie's release account for 97 per cent of the movie's box office revenue (Liu, 2006; Sawhney and Eliashberg, 1996); therefore, we collected box office revenue data on a weekly basis for eight weeks after the release of each movie. We also collected aggregate box office revenue (in thousands of US dollars) for the entire period of a movie's run. Box office data were collected primarily from www.numbers.com and verified using other sources, such as the Variety, ERC Box Office and IMDb movie databases (data sources listed in Table I.

3.3.4 Other independent variables. Four other independent variables (apart from CE and expert evaluation variables) that had an influence on box office revenues were included in the regression analyses in Tables IV and V. The first variable was the "number of screens" on which each movie was released, and this variable which has been demonstrated in the past to have a strong influence on box office revenues (Swami *et al.*, 1999). The second variable was the "production budget", which has also been demonstrated to have an effect on box office revenues (Basuroy *et al.*, 2003). The third and fourth variables were the total "number of movie critic reviews"[5] and the total "number of member comments"[6]. The number of screens and the total number of member comments varied for each week after release, and hence both were controlled for each week during the eight-week post-release periods.

4. Analysis, interpretation and discussion of results

An analysis of the data collected from various sources was performed using statistical analysis techniques such as Pearson's correlation and multiple linear regressions. All the variables (Table I) used in these analyses are numerical variables (interval scales and ratio scales) thereby allowing us to carry out correlational and regression analyses. The PASW Statistics 18 software was used for these analyses, and the results are presented in the form of Tables II-V. We begin by presenting the correlation results with respect to expert (film reviewers) and CE in order to fulfil the first objective of this study (i.e. to verify whether CE are on par with expert evaluations). Next, we present the regression results in order to explore whether box office performance closely mirrored expert evaluations or CE.

4.1 CE: are they on par with expert evaluations?

A Pearson's correlation analysis was used to determine whether there was a significant association between film reviewer evaluations and CE. As community members

evaluate movies throughout their theatre run, the CE data from each successive week after release (for up to eight weeks) were individually correlated with the film reviewer evaluation data. Moreover, both cumulative values[7] and additional[8]/ successive values of CE data were used in the Pearson's correlation analysis and the findings are presented in separate tables (Tables II and III for cumulative values, and Tables IV and V for additional values).

Table II provides the results of the Pearson's correlation between community and film reviewer evaluations, as measured in terms of average ratings. At the cumulative level, CE during the fifth week were found to be significantly correlated with film reviewers' evaluations (at $p < 0.05$), although this correlation was not very strong (0.117). During all of the other weeks, as well as during the theatre lifetime (the last row

	Evaluation period	Film reviewers' evaluations (variable name: FRE_Average Ratings)
Community-evaluation (variable name: CE_Average Ratings)	Upto week1	0.037
	Upto week2	0.086
	Upto week3	0.116
	Upto week4	0.109
	Upto week5	0.117*
	Upto week6	0.094
	Upto week7	0.097
	Upto week8	0.067
	Lifetime	0.070

Table II.
Pearson correlations between community evaluation (using cumulative values) and film reviewers' evaluation

Note: *Correlation is significant at 0.05 level (two-tailed)

		Community-evaluation (variable name: CE_Average Ratings)							
	Evaluation period	Upto wk1	Upto wk2	Upto wk3	Upto wk4	Upto wk5	Upto wk6	Upto wk7	Upto wk8
Community-evaluation (variable name: CE_Average Ratings)	Upto week1								
	Upto week2	*0.944***							
	Upto week3		*0.993***						
	Upto week4			*0.996***					
	Upto week5				*0.998***				
	Upto week6					*0.983***			
	Upto week7						*0.999***		
	Upto week8							*0.999***	
	Lifetime	*0.922***	*0.980***	*0.990***	*0.994***	*0.981***	*0.998***	*0.999***	*1***

Table III.
Pearson correlations of community evaluation (using cumulative values) during successive weeks

Note: **Correlation is significant at 0.01 level (two-tailed)

in Table II), there was no significant association between film reviewer and CE. This finding confirms that community members evaluate movies differently than film reviewers.

One reason for this difference is that community members and film reviewers are exposed to different sets of conditions which influence the nature of their evaluations of new movies. On the one hand, community member evaluations are based on the level of enjoyment that they got from watching the movies. Therefore, their reviews reflect their subjective movie-watching experience and satisfaction level. On the other hand, film reviewers watch a new movie during special screenings or previews with the intention of writing reviews for the magazines, newspapers or broadcasters that they represent. Thus, various stakeholders may influence a film reviewer's reviewing behaviour and, moreover, the film reviewers themselves might be endogenously biased in their attempt to write reviews according to the interests of their readers or audience. This bias occurs because the reviewers become familiar with their audience's interests as they gain reviewing experience, which may influence their opinions and reviews (Eliashberg and Shugan, 1997). For example, if the majority of a reviewer's readers are teenagers who may prefer romantic or action movies, the reviewer may be inclined to review such movies more positively or at least overstate the positive elements of these movies. Moreover, reviewing movies is a film reviewer's job (hence, an obligation), whereas, for community members, reviewing movies may be a passion or an enjoyable activity (hence, a voluntary activity). Table III provides the inter-correlation of CE during successive weeks. For example, the correlation between CE during week 1 and week 2 is 0.944 (at $p < 0.01$), week 2 and 3 is 0.933 (at $p < 0.01$) and so on. This correlation is found to be consistent during each successive week throughout the lifetime of movies suggesting that CE is more or less decided during the first few weeks after a movie's release. This could be explained by the fact that the majority of moviegoers watch the movie during the initial weeks after a movie's theatre release.

Tables IV and V below display the results obtained using only "additional" or new members' evaluation data from successive weeks. These two levels of analysis (i.e. cumulative values (used in Tables II and III) or additional values (used in Tables IV and V)) were used to explore whether CE were adequate at only the cumulative level or whether they were also suitable when the evaluations were based on the small groups of additional members during the successive weeks. User reviews and comments from a given week are known to exhibit word-of-mouth (WOM) influence over moviegoers in subsequent weeks

	Evaluation period	Film reviewers' evaluations (variable name: FRE_Average Ratings)
Community-evaluation (variable name: CE_Average Ratings)	Upto week1	*0.174***
	Upto week2	*0.370***
	Upto week3	*0.313***
	Upto week4	*0.190***
	Upto week5	*0.219***
	Upto week6	*0.208***
	Upto week7	*0.149*
	Upto week8	*0.149*
	Lifetime	*0.070*

Note: **Correlation is significant at 0.01 level (two-tailed)

Table IV. Pearson correlations between community evaluation (using additional values) and film reviewers' evaluations

Table V.
Pearson correlations of community evaluation (using additional values) during successive weeks

		Community-evaluation (variable name: CE_Average Ratings)							
	Evaluation period	Upto wk1	Upto wk2	Upto wk3	Upto wk4	Upto wk5	Upto wk6	Upto wk7	Upto wk8
Community-evaluation (variable name: CE_Average Ratings)	Upto week1								
	Upto week2	0.479**							
	Upto week3		0.514**						
	Upto week4			0.425**					
	Upto week5				0.423**				
	Upto week6					0.603**			
	Upto week7						0.485**		
	Upto week8							0.402**	
	Lifetime	*0.767***	*0.445***	*0.432***	*0.585***	*0.498***	*0.687***	*0.515***	*0.464***

Note: **Correlation is significant at 0.01 level (two-tailed)

(Chakravarty *et al.*, 2010; Dellarocas *et al.*, 2007; Godes and Mayzlin, 2004; Liu, 2006). WOM "involves informal communication among consumers about products and services" (Liu, 2006, p. 74). Experiential products such as movies are difficult to evaluate prior to purchase. Here consumers are known to engage in WOM to access information and/or exchange opinions. For example, all member comments formulated before a movie's release will influence the opening week box office revenues, whereas the comments articulated during the opening week (by members who watched the movie during the first week) will influence moviegoers during the second week, and so on. Such WOM effects are more prominent at the additional level, which led to our decision to also represent the Pearson's correlation results at this level (Tables IV and V).

Table IV provides the correlation results obtained for film reviewers' and CE based on only additional or successive new users' reviews, and these results are thus different from the results provided in Table II. In Table IV, community and film reviewer evaluations positively correlated ($p < 0.01$) during the first six weeks; however, this correlation was weak implying that during successive weeks, a small percentage of new group members have opinions similar to those of film reviewers. Nonetheless, the overall difference between community and film reviewers' evaluations remains. Table V provides the inter-correlation of CE (using only additional values of each week) for successive weeks. For example, the correlation between CE during week 1 and week 2 is 0.479 (at $p < 0.01$), week 2 and 3 is 0.514 (at $p < 0.01$) and so on. This correlation is found to be less consistent (almost half the coefficient compared to the values in Table II) during each successive week throughout the lifetime of movies, suggesting that community members who participate during successive weeks have different opinions compared to community members who evaluated the movies in the previous week. To summarise, from the results of our correlational analysis presented in Tables II-V we conclude that CE of new movies are different from those of film reviewers' evaluations.

As our findings show that CE are different from film reviewer evaluations, and the fact that both evaluations cannot be true simultaneously in such situations, our next goal was to establish which one of the two evaluations is correct or more accurate. As both film reviewers and community members are evaluating the quality of movies in terms of their enjoyment and satisfaction values, and the fact that high-quality movies should lead to higher box office revenues, the correct evaluation will be the one which better predicts box office revenues. Thus, in order to understand whether film reviewer evaluations or CE of the quality of movies are true, we compare the predictive power of CE and expert evaluations on market performance (box office revenues) in the following section.

4.2 Which group makes more reliable and accurate evaluations: film reviewers or online communities?

Average ratings of movies found in online communities are dynamic in nature and they are subject to change if future moviegoers give a different rating compared to those given by earlier moviegoers. However, film reviewers' evaluations are rather static since they rate a movie only once and they give their ratings in the first week of a movie's release. Hence, movie critic evaluations are more likely to influence moviegoers only in the opening week, whereas, online CE are more likely to influence moviegoers not only in the opening week but also throughout the lifetime of a movie. With the availability of several movie online communities, potential moviegoers are more likely to read the ratings and reviews posted by their peers, and hence future moviegoers consult online CE to decide whether to watch a movie or not. A 2010 research study by

Stradella Road (2010) shows that 52 per cent of moviegoers (in the USA and Canada) learn details about movies from the internet. More specifically, moviegoers get "review information" mostly from the internet (62 per cent from internet, 51 per cent from friends, 49 per cent on TV, 29 per cent from newspapers, 21 per cent from magazines, 19 per cent from radio, 6 per cent from elsewhere). Thus, the internet is the most sought after for movie review information compared to all other sources of information. With the ever-increasing internet penetration globally and the continuous mushrooming of new online communities, the internet and specifically online communities play a very important role as useful sources of information which influences consumer buying decisions.

Irrespective of whether online CE reflect customer satisfaction (of those moviegoers who have seen the movie) or expectation levels (of those future moviegoers who have not yet seen the movie), future moviegoers look to online CE as a reliable source of information more than movie critic evaluations. Moreover, future moviegoers, who check this information, do not really care about or do not really know whether the CE reflect customer satisfaction or customer expectation levels. The argument that online CE is better than film reviewer evaluation is supported by our findings which are presented in Tables II and III. However, to further validate this argument, we did a regression analysis (Tables IV and V) using box office revenues (market performance) as the dependent variable, and using online CE and movie critic evaluation as the main independent variables along with other variables that influence box office revenues. If online CE is better than movie critic evaluation, it means that future moviegoers are influenced more by online CE than movie critic evaluation, and if that is the case then the predictive power of online CE must be higher than movie critic evaluation.

If a movie is accepted by the majority of moviegoers, this acceptance will be reflected in the form of higher box office revenues. Thus, we regressed both film reviewer and CE using weekly box office revenues as dependent variables. To avoid multi-collinearity, the percentage of positive reviews in terms of CE and film reviewer evaluations (Table VI) were regressed separately with box office revenues (dependent variable) and then with the percentage of negative reviews (Table VII).

Film producers provide special screenings for film reviewers hoping to generate positive reviews for their movies. The influence of these reviews is expected to be high especially during opening week where the reviews are available to readers days ahead of the movie release (Eliashberg and Shugan, 1997). However, the results reported here (Table VI) indicate that film reviewers' evaluations have a weak effect on opening-week box office revenues (standardised coefficient $\beta = 0.075$, $p < 0.05$). The film reviewers' evaluations were positive and significant with regard to box office revenues in later weeks, particularly during week 2 ($p < 0.1$), week 3 ($p < 0.05$), week 6 and week 8 ($p < 0.005$); however, this relationship was weak. Conversely, the CE were positive and significant with regard to opening week box office revenues (standardised coefficient $\beta = 0.221$, $p < 0.001$). This result implies that CE have more predictive power than film reviewers' evaluations during the opening week. CE were positive and significant ($p < 0.05$) even during later weeks, particularly weeks 2, 3 and 4; however, this relationship was not strong. From the fifth until the eighth week after release, the CE were insignificant. This result might have occurred because more members had watched the movie by this point, resulting in a weaker WOM influence in later weeks than during the opening week.

In the previous section, we argued that the differences between community and film reviewers' evaluations may stem from their exposure to different sets of conditions. If we

Table VI.
Regression results: positive evaluations of film reviewers vs online community on box-office revenues

Independent variables	Weeks after release 1	2	3	4	5	6	7	8	Gross
Film reviewers' evaluations[a,c]	*0.075*	0.081	0.074	0.068	0.081	0.171	−0.051	0.176	*0.029*
t-statistic (*p*-value)	*2.294 (0.025)*	2.021 (0.046)	1.799 (0.076)	1.394 (0.168)	1.561 (0.123)	3.093 (0.003)	−0.573 (0.569)	3.181 (0.003)	*1.030 (0.307)*
Community evaluations[b,c]	*0.221*	0.083	0.098	0.104	0.054	0.024	0.045	−0.03	*0.052*
t-statistic (*p*-value)	*4.955 (0.000)*	2.455 (0.016)	2.637 (0.010)	2.397 (0.019)	1.239 (0.220)	0.477 (0.635)	0.587 (0.560)	−0.640 (0.526)	*1.688 (0.099)*
FRE total number of reviews[c]	0.024	0.134	0.097	0.057	0.038	0.023	0.183	−0.024	*0.050*
t-statistic (*p*-value)	0.518 (0.606)	2.977 (0.004)	2.267 (0.026)	1.155 (0.252)	0.715 (0.477)	0.389 (0.699)	1.982 (0.053)	−0.403 (0.689)	*1.422 (0.159)*
CE_Total number of reviews[c]	0.06	0.059	0.259	0.223	0.258	0.127	0.171	0.129	*0.216*
t-statistic (*p*-value)	2.124 (0.037)	1.670 (0.099)	6.260 (0.000)	4.511 (0.000)	4.947 (0.000)	1.879 (0.065)	1.468 (0.148)	1.870 (0.068)	*4.073 (0.000)*
Number of screens[c]	0.637	0.823	0.752	0.784	0.723	0.811	0.665	0.791	*0.701*
t-statistic (*p*-value)	13.964 (0.000)	2.647 (0.000)	16.758 (0.000)	15.674 (0.000)	14.003 (0.000)	11.217 (0.000)	5.796 (0.000)	11.113 (0.000)	*12.958 (0.000)*
Production budget[c]	0.221	0.064	−0.061	−0.018	0.137	0.02	0.066	0.09	*0.029*
t-statistic (*p*-value)	4.955 (0.000)	1.553 (0.124)	−1.406 (0.163)	−0.387 (0.700)	2.564 (0.013)	0.291 (0.772)	0.620 (0.538)	1.693 (0.098)	*0.928 (0.357)*
F-ratio (*p*-value)	221.079 (0.000)	159.891 (0.000)	149.094 (0.000)	104.578 (0.000)	92.302 (0.000)	73.203 (0.000)	21.677 (0.000)	75.190 (0.000)	*291.405 (0.000)*
Multiple R^2 (adjusted R^2)	0.948 (0.944)	0.919 (0.914)	0.918 (0.912)	0.896 (0.887)	0.891 (0.881)	0.883 (0.871)	0.722 (0.689)	0.913 (0.901)	*0.959 (0.956)*

Notes: Dependent variable: box-office revenues (weekly (i.e. from week 1 to week 8) and gross). [a]Variable name: FRE_Percentage Negative Reviews; [b]variable name: CE_Percentage Negative Reviews; [c]standardised coefficient (β)

Table VII.
Regression results: negative evaluations of film reviewers vs online community on box-office revenues

Independent variables	Weeks after release								
	1	2	3	4	5	6	7	8	Gross
Film reviewers' evaluations[a,c]	−0.032	−0.035	−0.064	−0.106	−0.032	−0.105	0.158	−0.138	*−0.005*
t-statistic (*p*-value)	−0.005 (0.323)	−0.946 (0.347)	−1.712 (0.091)	−2.362 (0.021)	−0.647 (0.520)	−1.941 (0.057)	1.969 (0.054)	−2.316 (0.025)	*0.183 (0.856)*
Community evaluations[b,c]	−0.118	−0.065	−0.087	−0.08	−0.009	−0.073	0.021	0.033	*0.000*
t-statistic (*p*-value)	−2.824 (0.006)	−1.924 (0.058)	−2.388 (0.019)	−1.959 (0.054)	−0.200 (0.842)	−1.445 (0.154)	0.283 (0.778)	0.658 (0.514)	*0.003 (0.998)*
FRE total number of reviews[c]	0.104	0.179	0.113	0.041	0.075	0.067	0.227	−0.004	*0.069*
t-statistic (*p*-value)	2.245 (0.028)	4.204 (0.000)	2.661 (0.009)	0.837 (0.405)	1.397 (0.167)	1.131 (0.263)	2.639 (0.011)	−0.059 (0.953)	*2.072 (0.042)*
CE_Total number of reviews [c]	0.044	0.061	0.273	0.247	0.271	0.126	0.174	0.145	*0.238*
t-statistic (*p*-value)	1.395 (0.167)	1.678 (0.097)	6.503 (0.000)	5.242 (0.000)	5.058 (0.000)	1.811 (0.075)	1.541 (0.130)	1.980 (0.054)	*4.19 (0.000)*
Number of screens[c]	0.678	0.811	0.737	0.761	0.725	0.83	0.67	0.79	*0.704*
t-statistic (*p*-value)	14.966 (0.000)	2.030 (0.000)	16.274 (0.000)	15.855 (0.000)	13.527 (0.000)	11.197 (0.000)	6.032 (0.000)	1.440 (0.000)	*12.586 (0.000)*
Production budget[c]	0.188	0.047	−0.073	−0.013	0.131	−0.026	0.07	0.071	*0.027*
t-statistic (*p*-value)	3.624 (0.001)	1.091 (0.278)	−1.657 (0.102)	−0.274 (0.784)	2.300 (0.025)	−0.358 (0.722)	0.690 (0.493)	1.276 (0.209)	*0.848 (0.339)*
F-ratio (*p*-value)	173.037 (0.000)	145.117 (0.000)	139.982 (0.000)	106.468 (0.000)	85.109 (0.000)	67.575 (0.000)	23.788 (0.000)	67.639 (0.000)	*276.065 (0.000)*
Multiple R^2 (adjusted R^2)	0.934 (0.929)	0.912 (0.906)	0.913 (0.907)	0.897 (0.889)	0.882 (0.872)	0.875 (0.862)	0.741 (0.709)	0.904 (0.891)	*0.957 (0.953)*

Notes: Dependent variable: box-ofice revenues (weekly (i.e. from week 1 to week 8) and gross). [a]Variable name: FRE_Percentage Negative Reviews; [b]variable name: CE_Percentage Negative Reviews; [c]standardised coefficient (β)

further explore this line of reasoning, it may also facilitate our explanation as to why CE more closely mirror market performance. Film reviewers are unlikely to possess equal levels of expertise in all movie genres, and their reviews are also likely to be influenced by various stakeholders. Thus, even if reviewers like (or dislike) some movies (or the actors, characters and so on in movies), they might provide a different evaluation to please their readers or audience. To achieve this goal, it is crucial that film reviewers understand the preferences or tastes of their readers. However, the proportion of their readers to the overall market of moviegoers is likely to be very small. Thus, their expertise is at risk of becoming narrower over time. In addition to this issue, it is nearly impossible for film reviewers to maintain 100 per cent accuracy. Additionally, other factors, such as the cast (star power), the movie director and genre, may overwhelm film reviewers' influence on box office revenues (Basuroy *et al.*, 2003; Liu, 2006).

Conversely, although community members' evaluations are subjective in nature, when they were "summed up", we gained insight into the community's "overall" opinion regarding each movie (i.e. community's majority opinion = total positive individual evaluations – total negative evaluations). Therefore, although community members' evaluations reflect their individual, subjective experiences, their collective evaluations are less likely to be "biased" and more likely to accurately reflect market behaviour (i.e. community behaviour mirrors overall market behaviour).

Another reason for this mirroring may be that each movie is evaluated by a different set of community members, who voluntarily choose to participate and write reviews, and such self-selection may depend on movie characteristics, individual preferences and tastes. Therefore, community members with specific interests in a movie genre may write reviews for those movies that align with their interests. Different movies may thus receive reviews from different sets of community members, and such evaluations are likely to be more "valid" and "reliable" than reviews written by the same individual. This difference in validity and reliability exists due to the improbability of an individual being an expert in all genres of movies but rather in one or a few genre. Unlike this, film reviewers have an obligation to write reviews of all movies that are released and because an average of eight movies is released each Friday in US theatres, their evaluations may not always be accurate for all types of movies and on all occasions.

An additional reason may be the awareness produced by positive WOM. More specifically, the higher community members' satisfaction level with a movie, the more people are likely to watch the movie in subsequent weeks. This WOM effect is confirmed by the strong positive correlation between CE obtained for adjacent weeks during the eight-week study period.

In other words, the CE variables used in this study were the net outcome of the effects of the individual member characteristics (i.e. preferences, tastes and past movie-watching experiences); the movie characteristics (i.e. major stars, genre, storyline and directors); and the social influences of advertising (trailers), community members, family, friends and WOM. Film reviewers do not consider all of these individual moviegoer effects, whereas overall CE do consider them, meaning that the latter are more likely to accurately reflect all moviegoers (the mass market). This explains why box office revenues, both on a weekly basis and at the gross level, mirror CE more closely than expert/film reviewer evaluations.

The results also demonstrate that the independent variable "CE_Total number of reviews" had a positive and significant relationship with box office revenues for nearly the entire movie life span (except for week 7). This is because the variable

"CE_Total number of reviews" or the total number of comments posted by community members is a satisfactory measure of the volume of WOM; according to Liu (2006), the volume of WOM has a significant effect on box office revenues by spreading more awareness about movies throughout the community. The next independent variable, "FRE_Total number of reviews", had a positive (but weak) relationship with box office revenues, with $p < 0.05$ during weeks 2, 3 and 7. As expected, the third independent variable, "number of screens", had a strong, positive and significant relationship with box office revenues, with $p < 0.001$ for all eight weeks as well as at the gross level. This finding suggests that the higher the number of screens showing a particular movie, the higher the box office revenues will be. The last independent variable, "production budget", was significant during the opening week ($p < 0.001$) and in the fifth week ($p < 0.05$), suggesting that a large movie budget increases public awareness of that movie. However, in later weeks, there is substantially more influence from WOM based on the actual movie-watching experience, rendering the effects of the production budget insignificant during this time.

Table VII below provides the regression results obtained using the percentage of film reviewers' and community members' negative evaluations on box office revenues. CE exhibited a significant but negative relationship with box office revenues during the first four weeks, whereas film reviewers' evaluations were not significant during the first two weeks. The relationship was negative because with an increase in the percentage of negative reviews, it created a negative WOM effect, which eventually decreased the box office revenue. As with the previous analysis in Table VI, the total number of community reviews had a positive and significant influence on box office revenues for nearly the entire movie life span (except week 1 and week 7).

As already mentioned, our findings in Tables IV and V shows that online CE have a strong positive and significant influence on box office revenues in the opening week as well as for gross box office revenues compared to film reviewers' evaluation. The findings also show that online CE is significant for almost the first four weeks following a movie's release, after that it becomes insignificant. This insignificance in later weeks could partially be explained by the fact that by this time other new movies have been released in the theatre (on average eight to ten new movies are released in theatres each week). Our findings, which show that online CE can predict market performance, are in line with the findings by Divakaran (2012) who showed that online community adoption behaviour closely represents market adoption behaviour, and that an online community is a good representation of its off-line market.

To summarise, the results from Tables IV and V support our original hypothesis that an online community meeting the three conditions (identified in Section 2.3) for reliable performance of evaluation tasks, can produce accurate collective evaluations. Having met these conditions, the Fandango online community's collective evaluation of new movies is found to accurately mirror actual box office revenues.

5. Conclusions, implications, limitations and further research

This study compares evaluations provided by film reviewers and a movie online community. Our empirical findings demonstrate that CE do not correlate significantly with those of film reviewers suggesting that film reviewers and the communities evaluate movies differently. The results confirm that box office revenues reflect CE more closely than expert evaluations especially during the opening week (during which

film reviewers' evaluations were found to be insignificant). This result is surprising because the perception among movie producers is that film reviewers influence moviegoers particularly on a movie's opening weekend, which is why producers provide special screenings for film reviewers. The findings also support the assumption that if an online community meets the conditions of diversity and heterogeneity of community members; provision of aggregation mechanism in the community; and having an independent governance structure, it can provide reliable and accurate movie evaluations.

While an online CE is shown to have an effect on box office revenues in this study, it may well be argued that each week's box office revenues can actually influence CE in subsequent weeks (bandwagon effect). For example, Moon *et al.* (2010) have shown that movies with high early box office revenues enhance movie ratings in subsequent weeks, and they have also shown that high advertising spending on movies supported by high ratings maximises a movie's box office revenues. For a moviegoer, the evaluation provided by an online community is more useful for deciding whether to watch it or not than looking at the box office sales for that movie. This is because sales is just a quantitative number whereas online CE provides both qualitative information about the movie in the form of online ratings and reviews from other moviegoers in addition to quantitative information such as number of members who rated positively and negatively, etc.

This paper is not about building a forecasting model for predicting market performance; instead we used market performance as the dependent variable only to check whether online CE is better than movie critic evaluations or not. This is also the reason why the market performance variable is not used in the analysis reported in Tables II-V. Moreover, because we used market performance (box office revenues) as the dependent variable in Tables IV and V, we had to introduce other commonly used independent variables such as production budget, number of screens, etc., which are found in previous studies related to movie box office sales (even movie critic evaluation is used as an independent variable in prior studies).

The present study has some managerial implications. Because CE have more predictive power than those of film reviewers during the opening week, studios can use online communities for evaluations of new movies. For example, in the case of movies for which the existing awareness level is low as reflected by less number of community members participation, film promotion managers can develop better trailers, alter the frequency of advertising, etc., to increase awareness level. If the CE is positive in the opening weeks, then film producers can negotiate with distributors and theatre owners regarding the number of screens and number of days their movies should run, etc. Thus, it would be wise for movie producers and studios to direct their focus towards online communities instead of film reviewers and attract more members to visit online communities and participate in online activities, which would further improve the reliability of new movie evaluations. Unlike film critics whose reviews appear mostly in the opening week of a movie's release, community members' reviews appear throughout the life of a movie in the theatre, and our study have shown that CE data from the opening week till fourth week could predict market sales. Hence movie makers must pay attention to what is spoken in such communities throughout its theatrical life time to better understand consumers' attitude, perceptions, opinions, etc., towards their movie. This study also highlights that future movie-goers rely more on the evaluation by their peers (i.e. community evaluation) than film critics' evaluation for deciding their movie choice, and hence film producers

should give more priority to movie-based online communities and provide incentives such as discounts for increasing community participation (Fandango already does that). Higher the number of community members participating in online discussions, better the community's prediction of new movies be.

Capitalising on the predictive potential of online communities could turn out to be vital in a time when so many industries are increasingly consumer driven. Movie studios are gradually realising this phenomenon. For instance, director Sam Raimi (the creator of superhero movies such as the original Spiderman trilogy and Darkman), who is currently making a movie based on the critically acclaimed videogame "The Last of Us", is making it a priority to acknowledge the demands of the zealous fans. Recently the director thus made a surprise appearance at the San Diego ComicCon 2014, announcing that the movie is under way (*MStars News*, 2014). As a way of ensuring that he will not betray the fans' vision of this iconic videogame, the director has relinquished more creative control to the original creator of the videogame than it has ever been the case in a movie adaptation of a videogame (Uproxx, 2014). This move seems to have created hype in online videogame communities. Although it is too early to speculate about the effects of such strategy, this example indicates that it is essential to obtain a deeper understanding of online communities, movie factors and movie performance. Indeed, a creative partnership between fans and the movie crew of the 2006 movie "Snakes on a Plane" emerged after online community members and bloggers created their own movie trailers, movie posters, songs and poems inspired by this movie – before the movie's release. The director and the studio behind the movie were on their toes and decided to re-shoot parts of the movie to include the ideas of fans (*New York Times*, 2001). Although the movie was not a box office success, this emergence of community relationships and their impact on movie production and performance raises the need for research on the direct and indirect influence of communities on movie performance.

Some limitations apply to the use and interpretation of the results of this study. Fandango is a US-based online community; hence the findings are limited to this geographic context. Indeed, Lee (2009) found that cultural differences influence the box office performance of movies. Lee (2009) investigated the effect of drama awards (i.e. best director award, best leading and supporting actor/actress, best screenplay and best movie editing) and non-drama awards (all other awards) on box office performance. He found that movies with American storylines as well as movies that contain relatively more special effects do not experience the same level of success in East Asia as in the USA (Lee, 2009). Also, the more culturally distant countries were from the USA, the more pronounced the negative effect of drama awards and East Asian box office. Further research is therefore needed to understand whether our findings extend to other regions.

Moreover, any online community or social network may be subject to manipulation by various agencies that may have a stake or interest, e.g. studio marketing managers. Nevertheless, the independent governance structure of online communities helps prevent such manipulations and red flag such members. Thus, further research could examine the prevalence and impact of such manipulation. The size of online communities such as Fandango could be considered an important impact factor in terms of the effectiveness of communities as forecasting tools and influencers of product market performance. Consequently, further research might examine whether our findings could be extended to those industries that do not

attract user participation in online communities to the same extent as the movie industry. In our study, data were collected and analysed at the community and movie level without differentiating between member segments. Therefore, future studies could also consider dividing community members into relevant groups, such as frequent and infrequent community participants, or frequent and infrequent moviegoers. Community members participate in online discussions not just after a movie's release but long before that, or as soon as an announcement of an upcoming movie is made. Hence, future studies can explore whether such pre-release data can be used in predicting future market performance.

Notes

1. Community level refers to only those members who are part of the selected online community (Fandango). It does not refer to all moviegoers worldwide.
2. The terms film reviewers and film critics are used interchangeably in this study and they mean one and the same thing.
3. In this study, we use the term "film reviewers" or "film critics" instead of journalistic reviewers for the sake of convenience. However, they all carry the same meaning.
4. Theatrical Market Statistics. Retrieved from http://www.mpaa.org
5. The number of movie critic reviews was defined as the number of reviews available for each movie that were written by movie reviewers from different journals.
6. The number of member comments was defined as the number of user comments for each movie that were posted by members of the online community Fandango.
7. A cumulative value is defined as the average of all the ratings given by all the community members who participated from the time when a movie is released (i.e. from week zero) to another point in time. For example, a "cumulative" average rating during week two is obtained by taking the average of all the ratings given by all the community members who participated from week zero till the end of week two.
8. In contrast to "cumulative" values, an additional value is defined as the average of the ratings given by all the community members who participated only during a specific week. For example, an "additional" average rating for week two is obtained by taking the average of all the ratings given by the community members who participated from the end of week one till the end of week two.

References

Abony, G. (1983), "Filtering: an approach to generating the information base for collective choice", *Management Science*, Vol. 29 No. 4, pp. 409-418.

Adams, M. (2006), "Cutting out the critics", *Marketing Magazine*, 6 May, Vol. 111 No. 20, p. 19.

Basuroy, S., Chatterjee, S. and Ravid, S.A. (2003), "How critical are critical reviews? The box office effects of film critics, star power, and budgets", *Journal of Marketing*, Vol. 67 No. 4, pp. 103-117.

Boatwright, P., Basuroy, S. and Kamakura, W. (2007), "Reviewing the reviewers: the impact of individual film critics on box office performance", *Quantitative Marketing & Economics*, Vol. 5 No. 4, pp. 401-425.

Boor, M. (1992), "Relationships among ratings of motion pictures by viewers and six professional movie critics", *Psychological Reports*, Vol. 70 No. 3c, pp. 1011-1021.

Burnett, G. (2000), "Information exchange in virtual communities: a typology", *Information Research*, Vol. 5 No. 4, available at: http://informationr.net/ir/5-4/paper82.html

Carroll, J.M. and Rosson, M.M. (1996), "Theorizing mobility in community networks", *International Journal of Human-Computer Studies*, Vol. 66 No. 12, pp. 944-962.

Chakravarty, A., Young, L. and Mazumdar, T. (2010), "The differential effects of online word-of-mouth and critics' reviews on pre-release movie evaluation", *Journal of Interactive Marketing*, Vol. 24 No. 3, pp. 185-197.

D'Astous, A. and Touil, N. (1999), "Consumer evaluations of movies on the basis of critics' judgments", *Psychology & Marketing*, Vol. 16 No. 8, pp. 677-694.

Dellarocas, C., Xiaoquan, Z. and Awad, N.F. (2007), "Exploring the value of online product reviews in forecasting sales: the case of motion pictures", *Journal of Interactive Marketing*, Vol. 21 No. 4, pp. 23-45.

Denning, P.J. and Hayes-Roth, R. (2006), "Decision making in very large networks", *Communications of the ACM*, Vol. 49 No. 11, pp. 19-23.

Divakaran, P.K.P. (2012), "Pre-release member participation as potential predictors of post-release community members' adoption behaviour: evidence from the motion picture industry", *Behaviour & Information Technology*, Vol. 32 No. 6, pp. 545-559.

Duan, W., Gu, B. and Whinston, A.B. (2008), "Do online reviews matter? An empirical investigation of panel data", *Decision Support Systems*, Vol. 45 No. 4, pp. 1007-1016.

Eliashberg, J. and Shugan, S.M. (1997), "Film critics: influencers or predictors?", *Journal of Marketing*, Vol. 61 No. 2, pp. 68-78.

Eliashberg, J., Elberse, A. and Leenders, M.A.A.M. (2006), "The motion picture industry: critical issues in practice, current research, and new research directions", *Marketing Science*, Vol. 25 No. 6, pp. 638-661.

Eliashberg, J., Hui, S.K. and Zhang, Z.J. (2007), "From story line to box office: a new approach for green-lighting movie scripts", *Management Science*, Vol. 53 No. 6, pp. 881-893.

Franke, N. and Piller, F. (2004), "Value creation by toolkits for user innovation and design. The case of the watch market", *Journal of Product Innovation Management*, Vol. 21 No. 6, pp. 401-415.

Franke, N. and von Hippel, E. (2003), "Satisfying heterogeneous user needs via innovation toolkits: the case of Apache security software", *Research Policy*, Vol. 32 No. 7, pp. 1199-1215.

Gemser, G., Van Oostrum, M. and Leenders, M. (2007), "The impact of film reviews on the box office performance of art house versus mainstream motion pictures", *Journal of Cultural Economics*, Vol. 31 No. 1, pp. 43-63.

Godes, D. and Mayzlin, D. (2004), "Using online conversations to study word-of-mouth communication", *Marketing Science*, Vol. 23 No. 4, pp. 545-560.

Goldenberg, D., Nichols, D., Oki, B.M. and Terry, D. (1992), "Using collaborative filtering to weave an information tapestry", *Communications of the ACM*, Vol. 35 No. 12, pp. 61-70.

Hagel, J. III and Armstrong, A. (1997), *Net Gain: Expanding Markets Through Virtual Communities*, Harvard Business School Press Books, Boston, MA.

Hara, N. and Hew, K.F. (2007), "Knowledge-sharing in an online community of health-care professionals", *Information Technology & People*, Vol. 20 No. 3, pp. 235-261.

Holbrook, M.B. (1999), "Popular appeal versus expert judgments of motion pictures", *Journal of Consumer Research*, Vol. 26 No. 2, pp. 144-155.

Hsu, G. (2006), "Evaluative schemas and the attention of critics in the US film industry", *Industrial & Corporate Change*, Vol. 15 No. 3, pp. 467-496.

Iandoli, L., Klein, M. and Zollo, G. (2007), "Can we exploit collective intelligence for collaborative deliberation? The case of the climate change collaboratorium (12/2007)", MIT Sloan Research Paper No. 4675-08, available at: http://ssrn.com/abstract=1084069 (accessed 23 January 2009).

Jabeur, K. and Martel, J.-M. (2007), "A collective choice method based on individual preferences relational systems (p.r.s.)", *European Journal of Operational Research*, Vol. 177 No. 3, pp. 1549-1565.

Johnson, B.D. (2003), "Confessions of a critic", *Maclean's*, Vol. 116 No. 36, pp. 58-59.

Kim, H.-N., Ji, A.-T., Ha, I. and Jo, G.-S. (2010), "Collaborative filtering based on collaborative tagging for enhancing the quality of recommendation", *Electronic Commerce Research & Applications*, Vol. 9 No. 1, pp. 73-83.

Kobayashi, T. (2010), "Bridging social capital in online communities: heterogeneity and social tolerance of online game players in Japan", *Human Communication Research*, Vol. 36 No. 4, pp. 546-569.

Kozinets, R.V. (1999), "E-tribalized marketing?: the strategic implications of virtual communities of consumption", *European Management Journal*, Vol. 17 No. 3, pp. 252-264.

Kozinets, R.V. (2002), "The field behind the screen: using netnography for marketing research in online communities", *Journal of Marketing Research*, Vol. 39 No. 1, pp. 61-72.

Lee, F.L.F. (2009), "Cultural discount of cinematic achievement: the academy awards and US movies' East Asian box office", *Journal of Cultural Economics*, Vol. 33 No. 4, pp. 239-263.

Leimeister, J.M., Schweizer, K., Leimeister, S. and Krcmar, H. (2008), "Do virtual communities matter for the social support of patients?: antecedents and effects of virtual relationships in online communities", *Information Technology & People*, Vol. 21 No. 4, pp. 350-374.

Lin, H.-F. and Lee, G.-G. (2006), "Determinants of success for online communities: an empirical study", *Behaviour & Information Technology*, Vol. 25 No. 6, pp. 479-488.

Liu, Y. (2006), "Word of mouth for movies: its dynamics and impact on box office revenue", *Journal of Marketing*, Vol. 70 No. 3, pp. 74-89.

Luo, G.Y. (2002), "Collective decision-making and heterogeneity in tastes", *Journal of Business & Economic Statistics*, Vol. 20 No. 2, pp. 213-226.

Madupu, V. and Cooley, D.O. (2010), "Antecedents and consequences of online brand community participation: a conceptual framework", *Journal of Internet Commerce*, Vol. 9 No. 2, pp. 127-147.

Moon, S., Bergey, P.K. and Iacobucci, D. (2010), "Dynamic effects among movie ratings, movie revenues, and viewer satisfaction", *Journal of Marketing*, Vol. 74 No. 1, pp. 108-121.

MStars News (2014), "ComicCon 2014 announcement: Sam Raimi talks 'The Last Of Us' movie & 'Evil Dead' TV show, by Solis, J.", *MStars News*, 25 July, available at: www.mstarz.com/articles/34304/20140725/comic-con-2014-announcement-sam-raimi-talks-the-last-of-us-movie-evil-dead-tv-show.htm (accessed 15 December 2014).

Nambisan, S. (2002), "Designing virtual customer environments for new product development: toward a theory", *Academy of Management Review*, Vol. 27 No. 3, pp. 392-413.

Nambisan, S. and Baron, R.A. (2010), "Different roles, different strokes: organizing virtual customer environments to promote two types of customer contributions", *Organization Science*, Vol. 21 No. 2, pp. 554-572.

New York Times (2001), "After hype online, 'Snakes on a Plane' is let down at box office", by Waxman, S., *New York Times*, 21 August, available at: www.nytimes.com/2006/08/21/movies/21box.html?_r = 0

Nolan, T., Brizland, R. and Macaulay, L. (2007), "Individual trust and development of online business communities", *Information Technology & People*, Vol. 20 No. 1, pp. 53-71.

Oral, M., Kettani, O. and Lang, P. (1991), "A methodology for collective evaluation and selection of industrial R&D projects", *Management Science*, Vol. 37 No. 7, pp. 871-885.

Phang, C.W., Kankanhalli, A. and Sabherwal, R. (2009), "Usability and sociability in online communities: a comparative study of knowledge seeking and contribution", *Journal of the Association for Information Systems*, Vol. 10 No. 10, pp. 721-747.

Pitta, D.A. and Fowler, D. (2005), "Online consumer communities and their value to new product developers", *Journal of Product & Brand Management*, Vol. 14 No. 5, pp. 283-291.

Plucker, J.A., Kaufman, J.C., Temple, J.S. and Qian, M. (2009), "Do experts and novices evaluate movies the same way?", *Psychology & Marketing*, Vol. 26 No. 5, pp. 470-478.

Ravid, S.A., Wald, J. and Basuroy, S. (2006), "Distributors and film critics: does it take two to Tango?", *Journal of Cultural Economics*, Vol. 30 No. 3, pp. 201-218.

Reinstein, D.A. and Snyder, C.M. (2005), "The influence of expert reviews on consumer demand for experience goods: a case study of movie critics", *The Journal of Industrial Economics*, Vol. 53 No. 1, pp. 27-51.

Rheingold, H. (1993), *The Virtual Community: Homesteading on the Electronic Frontier*, Addison-Wesley, London.

Satnam, A. (2008), *Collective Intelligence in Action*, Manning Publications Co, Greenwich, CT.

Sawhney, M.S. and Eliashberg, J. (1996), "A parsimonious model for forecasting gross box-office revenues of motion pictures", *Marketing Science*, Vol. 15 No. 2, pp. 113-131.

Shrum, W. (1991), "Critics and publics: cultural mediation in highbrow and popular performing arts", *American Journal of Sociology*, Vol. 97 No. 2, pp. 347-375.

Stradella Road (2010), "MOVIEGOERS 2010", by Paddison, G., Middleton, L.S., & Nielsen NRG, Hollywood, LA.

Sunstein, C.R. (2006), *Infotopia: How Many Minds Produce Knowledge*, Oxford University Press Incorporated, New York, NY.

Surowiecki, J. (2004), *Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations*, Doubleday Publishing, New York, NY.

Swami, S., Eliashberg, J. and Weinberg, C.B. (1999), "SilverScreener: a modeling approach to movie screens management", *Marketing Science*, Vol. 18 No. 3, pp. 352-372.

Turnball, P.W. and Meenaghan, A. (1980), "Diffusion of innovation and opinion leadership", *European Journal of Marketing*, Vol. 14 No. 1, pp. 3-33.

Uproxx (2014), "Sam Raimi is making a 'The Last Of Us' movie, and it might star Arya Stark", by Birch, N., available at: http://uproxx.com/gammasquad/2014/07/sam-raimi-the-last-of-us-movie-maisie-williams/ (accessed 14 December 2014).

Venkatraman, M.P. (1989), "Opinion leaders, adopters, and communicative adopters: a role analysis", *Psychology & Marketing*, Vol. 6 No. 1, pp. 51-68.

About the authors

Pradeep Kumar Ponnamma Divakaran is an Assistant Professor at the ESC Rennes School of Business in France. He pursued his PhD at the Aarhus University, Denmark. Prior to this he obtained his Master's Degree in E-Business and ICT for Management from the Politecnico di Torino, Italy and an MBA from the Cardiff University, UK. His research interests include IT-enabled social networks and their role in new product development, preannouncements effects in the Web 2.0 world, online brand communities, user-driven innovation and social networks

enabled innovation. He is a Member of EMAC, EURAM, AOM and user and open innovation community. Assistant Professor Pradeep Kumar Ponnamma Divakaran is the corresponding author and can be contacted at: pradeep.divakaran@esc-rennes.com

Sladjana Nørskov is a Post-Doc and Part-time Lecturer at the Aarhus University, Denmark. She pursued her PhD at the Aarhus University at their Management Department. Prior to this, Sladjana obtained her Master's Degree in Business Administration at the Aarhus University, Denmark. Her research areas include innovation management, community and organisational change.