

Theory and Evidence...

You Talkin' to Me?: Using Internet Buzz as an Early
Predictor of Movie Box Office

By

Anthony Versaci

An honors thesis submitted in partial fulfillment

of the requirements for the degree of

Bachelor of Science

Undergraduate College

Leonard N. Stern School of Business

New York University

May 2009

Professor Marti G. Subrahmanyam

Faculty Advisor

Professor Samuel Craig

Thesis Advisor

Abstract

In this study, I attempt to capture prerelease Internet buzz for movies through the use of variables like trailer views, message board comments, and votes of desire. By utilizing these buzz variables, my objective was to determine whether Internet buzz provides additional predictive information in terms of a film's box office revenues beyond a film's individual characteristics like genre, star power, budget, and rating. For the 62 films in my sample, I tracked their Internet buzz three weeks prior to each of their release dates. Then, through linear regression, I assessed the statistical significance of the buzz variables in predicting opening weekend box office gross. In my analysis, my findings suggest that three of my four buzz variables, those corresponding to interest and desire, are statistically significant and positively related to opening weekend box office. I also find that including the buzz measures considerably increases the explanatory power of the model. After incorporating user and critic ratings as measures corresponding to film quality, I find that neither is significantly related to opening weekend box office. Lastly, I provide an initial attempt at evaluating what factors may contribute to buzz and find that budget, categorization as a sequel, and the action genre are positively related to buzz.

1. Introduction

In 1925, E.K. Strong first developed the popular marketing acronym, AIDA (Awareness, Interest, Desire, Action), which he attributed to Elias St. Elmo Lewis in 1898. It has since been included in numerous marketing textbooks and utilized by salesman for years. Many scholars have built upon it or have provided their own variations, perhaps the most notable of which is what has become known as the 'Hierarchy of Effects Model' by Lavidge and Steiner (1961). This model contributes a couple more levels and was supported with a psychological model of behavior corresponding to cognition, affect and behavior. The underlying idea behind these two models as well as their variations was that consumers do not simply buy products impulsively, but instead go through a certain cognitive process before they make a purchase. The length and depth of this process may vary based on the product, but consumers nonetheless undergo some pre-purchase process in order to come to a decision. For advertisers and salesmen, the implication was that they could not move consumers directly to purchase. Instead, they had to first move them through a series of cognitive steps. Advertisers would have to initially make consumers aware of their product, pique their interest in it, induce desire for the product, and lastly drive the consumer towards a purchase. Advertisers, however, are not the only major influencer capable of moving consumers along these steps. Another perhaps more powerful influence is that of word of mouth.

Word of mouth is simply the spread of information from person to person. When consumers are seeking to buy a product, they do not simply wait for an advertiser to inform them of the product that meets their needs. On the other hand, they ask family, friends, neighbors, coworkers, etc. for suggestions. Even when they are not actively looking for information, fellow consumers are always willing to share information on products, both good and bad. Particularly

worrisome for advertisers about word of mouth is the fact that the message delivered concerning their product is out of their control, not to mention that the messenger is likely more trustworthy. Even further, since word of mouth was traditionally only person-to-person communication, it was rather difficult to track. The Internet, however, has greatly changed this dynamic.

By enabling consumers to share information on products at both levels and speeds never experienced before, the Internet has dramatically enhanced the diffusion of word of mouth. Consumers of all product categories are communicating with one another and sharing their opinions on blogs, message boards, chat rooms, and social-networking sites all over the Internet. Not only has the Internet enhanced the spread of both the consumers' messages as well as the advertisers' attempts to influence them, it has also provided interested parties with a convenient record of word of mouth, one not provided by person to person communication. This gives advertisers and firms the prime opportunity to track word of mouth and see exactly what consumers may be saying about their products. The movie industry is certainly no exception in this regard.

Prior to the advent of the Internet, the only way for movie studios or researchers to gauge consumer opinion, word-of-mouth, and the overall buzz for films was through surveys or focus groups. Now, with the widespread usage of the Internet, studios and researchers have another more comprehensive resource to measure buzz. For example, one tool researchers can currently use in post-release is user ratings. User ratings are used throughout the Internet for all varieties of products and allow users to provide instant feedback as well as to observe the general opinion of fellow consumers. In terms of movies, there are numerous Internet ratings sites including IMDb, Rotten Tomatoes, Netflix, Blockbuster Online, Fandango and Yahoo! Movies. However, since films are an experiential product, user ratings only can provide post-release information and thus

cannot be used as a measure of Internet buzz prior to a film's theatrical release. Gauging prerelease buzz, on the other hand, is arguably much more valuable to movie studios because of the significance of opening weekend box office.

First and foremost, opening weekend box office is crucial for a film because it sets the tone for the rest of the film's revenue windows, including the highly profitable DVD release. Opening weekend box office also makes up a major portion of a film's overall gross, many times even up to 50% (In 2003, movies received on average 41% of their total gross in their first week)¹, and if opening weekend sales disappoint, it is unlikely that a film will recover (Simonoff and Sparrow (2000) found a .93 correlation between logged opening weekend box office and logged total box office). Because of the extremely narrow timeframe studios have to witness either a success or failure and due to the lack of actionable strategies once a film is released, prerelease buzz is of the utmost importance. Prior to a film's release, some actions studios can take are to hire market research firms to gauge overall consumer awareness, interest and desire (Both Nielsen NRG² and MarketCast³ provide tracking data in this regard) and to conduct test screenings to determine what expectations may be, how well those expectations are met, and what word of mouth may be following release. The Internet, however, can provide additional sources for studios to assess consumer sentiment and awareness before opening weekend arrives.

Whether it is through blogs, message boards, reviews or some other means, consumers are rapidly sharing information on products with one another, and movies are no exception. There are countless websites, blogs, and message boards on the web devoted to movie news, reviews, discussion etc. Perhaps most chief of which is the Internet Movie Database

¹ Hayes, Dade and Jonathan Bing. *Open Wide: How Hollywood Box Office Became A National Obsession*. New York: Hyperion. 2004. pg. 8.

² en-us.nielsen.com/tab/industries/entertainment

³ www.marketcastonline.com

(www.imdb.com), a site with over 57 million visitors a month that, in addition to the 47 main boards, features a message board for each film on its site. Studios can now consult sites like IMDb to see firsthand what consumers are saying about their prospective products.

The Internet has not only enhanced the spread of word of mouth, but it has also changed the approach to a staple of movie promotion, the movie trailer. Movie trailers are short 1-2.5 minute montages, often accompanied by voiceover and music, which are used to generate awareness and excitement for an upcoming film. Before the Internet, movie trailers could only be seen in the theater during the coming attractions prior to the feature presentation. The trailers themselves are often costly to distribute and require studios to be strategic in selecting which films to place their trailers before in order to reach the appropriate target audience. This practice of course still goes on today, but the Internet has provided another venue for moviegoers to watch trailers. Now, typically coinciding with their theatrical release, trailers are released by studios on the Internet often to websites like Apple and Yahoo!. Soon after their initial release, these trailers can be found on blogs and message boards around the web and are often replicated as well on other video sites like YouTube. The Internet has thus provided studios with substantially more reach at a much lower cost. Additionally, studios could also potentially use the number of trailer views as a measure of interest, as I have done in this study.

The Internet has clearly affected some of the long held practices in the movie industry and has allowed studios and consumers alike to both provide and acquire information regarding films in ways they previously were not able. The key question, however, for studios is whether or not this new and easily accessible information provides them with any predictive value. This is the question that I aim to address to in my study. Under the AIDA framework, using popular film websites like Comingsoon.net (www.comingsoon.net) and TrailerAddict

(www.traileraddict.com), I tracked the number of message board comments and trailer views as a way of measuring interest. In addition, from the online movie ticket service, Fandango (www.fandango.com), I utilized its prerelease buzz polls that appear on each film's respective page, in order to create one variable for awareness and another for desire. For each film in my sample of 62, I tracked the respective values of each variable weekly for three weeks prior to each film's release date and on each's week of release. Utilizing these four variables as a way to encapsulate Internet buzz, I then used linear regression to determine whether or not Internet buzz provided predictive information over-and-above the predictive value provided by characteristics distinctive to the film like genre, star power, budget, and rating.

My paper is organized as follows: Following this introduction, the first section provides a brief discussion of some of the prior literature concerning box office forecasting, Internet buzz, and word of mouth in the movie industry. The second section provides a detailed description of the sample, the dependent variable and the independent variables, including the traditional, buzz and quality variables. The third section, divided into four subsections, details the results and provides a discussion of their implications. In order, the four subsections include the analysis and discussion of the traditional variables, the buzz variables, the quality variables, and lastly the question of what contributes to buzz. Finally, the paper ends with managerial implications and a discussion of potential further research.

2. Relevant Literature

Due to the challenge that predicting movie box office has and continues to present, there has been substantial research in this area. Much research has centered on evaluating how certain film characteristics contribute to box office. Litman (1983) conducted one of the earliest studies

in this regard and utilized a multiple regression model with production costs, genre, MPAA rating, presence of a star, major vs. independent release, Christmas release, and critics' ratings as independent variables. His study found evidence that major releases, Christmas releases, the science fiction genre, production costs, the presence of a star and critics' ratings were all significant predictors of total theatrical revenues.

More recently, Simonoff and Sparrow (2000) found that genre, MPAA rating, summer release and the presence of stars were all significant in relation to total box office. However, when they incorporated opening weekend box office and screens as independent variables, they found that MPAA rating and the presence of stars lose their significance. Terry et al. (2005) conducted a multiple regression analysis of total box office using many of the same independent variables noted above. Their study found that critical acclaim, Academy Award nominations, sequels, production budget and number of theaters were all statistically significant and positively associated with total box office. Unlike some of his peers, Ravid (1999) took a look at return in addition to revenues. Though he found that budget, volume of reviews, MPAA rating, and sequels were all significant in relation to revenues, he found that only ratings of G and PG, and with less significance, sequels and volume of reviews were significant in relation to returns.

With the advent of the Internet, however, researchers have now found another independent variable to use when predicting box office: word of mouth. While there has been considerable literature with regards to word of mouth and buzz in the movie industry as well as in the book and music industries, many of these studies have only considered online user ratings, a post-release measure. Dellarocas et al. (2007) analyzed user ratings from Yahoo! Movies and found that incorporating user ratings into their model provided significantly more accurate results in forecasting a film's total box office revenues. They found that volume (awareness

measured by total number of reviews), valence (consumer attitude measured by average user rating), and dispersion (spread of communication across communities measured by the distribution among users across gender and age) were all statistically significant predictors of future box office revenues. However, since this study's objective was to assess the predictive value of user ratings *from* opening weekend, no attempt was made to look at online word of mouth or buzz prior to release.

Earlier studies of online word of mouth have found results inconsistent with Dellarocas et al. (though their later study aimed to reconcile such inconsistencies). For example, Godes and Mayzlin (2004) assessed the explanatory power of online conversations with regards to TV ratings. They found that dispersion, and not volume, was statistically significant in providing such explanatory power. Even further, Duan et al. (2005), using Yahoo! Movie ratings, found that volume, but not valence, was significantly associated with higher movie revenues.

Likewise, Liu (2006) found similar results when he analyzed Yahoo! Movie message board posts. Classifying each as either positive, negative, mixed, neutral or irrelevant, he too found that volume, and not valence, provided statistically significant explanatory power for box office revenues. Of perhaps the most direct relevance to my study, however, is the fact that since Yahoo! Movie users posted on movies' message boards prior to their respective release dates, Liu also had a measure of prerelease online word of mouth and buzz. By looking at message board posts prior opening weekend, he found that prerelease buzz was also a significant indicator of total box office. It is in this area that my research aims to contribute. By tracking the number of prerelease trailer views, message board comments, and votes of desire, I have a variety of measures for prerelease volume of online word of mouth and buzz. Furthermore, by looking at

the percentages among the votes of desire, I also have a measure for prerelease valence of online word of mouth.

3. Data

3.1 The sample

The films included in my sample are all wide released films from November 7, 2008 through April 3, 2009. This window of time allowed me to capture data for the holiday movie season (marked unofficially by the release of *Madagascar: Escape 2 Africa* on November 7⁴) as well as Oscar season, the period in late December when films with perceived Oscar potential are released. By ending the data collection on April 3, I did stop short of the summer blockbuster season, which generally begins in May, but I was able to include major releases and summer blockbuster-esque films like *Watchmen*, *Monsters vs. Aliens*, and *Fast & Furious*. Also included in my sample are films that had platform releases with set wide release dates. For example, *Doubt* had a limited release on December 12 and a designated wide release on December 25 and, for the purposes of this analysis, December 25 was taken as its release date. On the other hand, films with more staggered releases like Best Picture winner *Slumdog Millionaire* were not included because one wide release date could not be pinpointed. Overall, the final sample consisted of 62 films.

3.2 Dependent Variable: Opening Weekend Box Office

Due to the limited timeframe of my sample, total box office gross could not be used as my dependent variable since films released earlier will simply have had more time to accumulate box office gross. Instead, I decided to use domestic opening weekend box office as my

⁴ McNary, Dave and Pamela McClintock. "High hoped for 'Madagascar' sequel." *Variety*. Nov. 6, 2008. <<http://www.variety.com/article/VR1117995403.html?categoryid=13&cs=>>

dependent variable. This was considered a viable substitute because of the large proportion of total box office gross that it constitutes. Opening weekend box office data was taken from Box Office Mojo (www.boxofficemojo.com). The film in my sample with the highest opening weekend box office was *Fast & Furious* while the film with lowest was *Delgo*.

Table 1: Opening Weekend Box Office Summary

Mean	20,720,651
Standard Error	2,221,542
Median	16,930,926
Standard Deviation	17,492,443
Minimum	511,920
Maximum	70,950,500

3.3 Independent Variables: Traditional Film Variables

3.3.1 Genre

Genre is a categorical variable and is thus evaluated with dummy variables. I determined a film’s genre by using the Internet Movie Database. Multiple genres are typically listed but, as per the site’s instructions, “the main genre should always be placed first.” Because of my relatively small sample, it was decided to define genre in five broad categories in order to limit degrees of freedom. These categories were Drama, Action, Comedy, Horror and Animated. For the most part, by using the first and second genres listed on IMDb, I was able to appropriately place each film under each of these categories. (*Jonas Brothers: The 3D Concert Experience*, a concert film, was loosely placed in the Comedy category due to its light tone.) A Comedy film like *Role Models* was thus assigned a value of 1 for the Comedy variable and a value of 0 for the other genre variables.

Table 2: Genre Breakdown

Genre	Number	Percent of Sample
Action	14	22.6%
Animated	6	9.7%
Comedy	20	32.3%
Drama	16	25.8%
Horror	6	9.7%

3.3.2 MPAA Rating

For the analysis, MPAA rating was treated as an ordinal variable. Accordingly, G = 1, PG = 2, PG-13 = 3, R = 4. For example, a film like *Role Models* with an R rating was given a Rating value of 4.

Table 3: MPAA Rating Breakdown

MPAA Rating	Number	Percent of Sample
G	2	3.2%
PG	15	24.2%
PG-13	28	45.2%
R	17	27.4%

3.3.3 Budget

Budget information for each film was found on Box Office Mojo. In cases that this information was not available there, IMDb's box office/business link was consulted. Finally, in the instances that budget was not available on either site, an average was taken from the sample based on genre and was then used as a proxy. James Bond film *Quantum of Solace* was the most expensive film in my sample with a budget of \$200 million while the relationship drama *Not Easily Broken* was the least expensive at \$5 million

Table 4: Budget Breakdown (Millions)

Mean	\$59.73
Standard Error	\$7.35
Median	\$39.00
Standard Deviation	\$48.78
Minimum	\$5.00
Maximum	\$200.00

3.3.4 Star Power

To assess a film's star power, I utilized the Forbes 2009 Star Currency list. Via a confidential survey of members of the entertainment industry, Forbes assembled a list of over 1400 actors and gave each a score from 0-10 based on his/her ability to contribute to financing, theatrical box office performance, and post-theatrical life. Among the criteria, of primary relevance to this analysis were the individual actor's ability to drive box office performance, his/her ability to attract audiences in any genre, and his/her popularity among most demographic groups⁵. Accordingly, a star like Will Smith was awarded the only 10.00 while a lesser known actress like Sasha Alexander was awarded a 0.66.

To provide a film with an overall star power score, I took the Star Currencies of the five highest rated stars of each film and added them together. By not using all of the actors included on the Forbes list for a given film, I was able to minimize the impact of lesser known stars as well as give major stars more weight. This approach also helped appropriately limit the star power of animated films that often feature a large number of stars voicing the characters. In my sample, the film with the highest Star Power score was the ensemble comedy *He's Just Not That*

⁵ "Star Currency: Complete Methodology." *Forbes.com*. Feb.10, 2009. <http://www.forbes.com/2009/02/06/forbes-star-currency-methodology-business-media-star-currency-09_0210_methodology.html>

Into You with a score of 36.76. The films with the lowest Star Power scores were *Miss March* and *Jonas Brothers 3D: The Concert Experience* with scores of 0.

Table 5: Star Power Breakdown

Mean	18.03
Standard Error	1.13
Median	18.07
Standard Deviation	8.87
Minimum	0.00
Maximum	36.76

3.3.5. Sequel

Whether a film was a sequel or not was also evaluated as a dummy variable. Sequels are generally put into production based on the success of a prior film and thus have a preexisting fan base. As a result, this is a significant factor to consider in the analysis. Films that are sequels received a value of 1 for the Sequel dummy variable and a 0 otherwise. For this analysis, I defined sequel more loosely in order to include reboots of film franchises like *Friday the 13th* (2009) and *Punisher: War Zone*. Studios generally produce reboots in order to reinvigorate declining franchises as a way to recapture the success of previous entries. Like traditional sequels (films that continue the narrative of a previous work), reboots also have built-in fan bases and, as such, I classified them as sequels in order to capture this characteristic. Remakes, on the other hand, were not classified as sequels because general audiences may not be familiar with the original film, especially if that film was a foreign language film or a film released many years ago. Of the 62 films in my sample, 9 were categorized as sequels (14.5%).

Table 6: Sequels Included in the Sample

Madagascar: Escape 2 Africa
Quantum of Solace
Transporter 3
Punisher: War Zone
Underworld: Rise of the Lycans
The Pink Panther 2
Friday the 13th (2009)
Madea Goes to Jail
Fast & Furious

3.4 Independent Variables: Internet Buzz Variables

3.4.1 Trailer Views

Before the prevalence of the Internet, the only place moviegoers could view trailers was at the theater during the coming attractions. Today, releasing a trailer on the Internet is now one of the prime ways studios can begin to promote their films and generate buzz. There are numerous websites where movie fans can view trailers (in varying quality) including Yahoo! Movies, Apple.com, and YouTube among many others. To gauge Internet buzz, I first attempted to use YouTube by weekly noting the number of views for the most watched trailer of each respective film. However, this did not prove feasible due to how frequently videos are taken down. Furthermore, because of the large number of uploaded trailers, it was also impractical to manually aggregate the views of all the trailers for each film

In the end, I decided to use TrailerAddict (www.traileraddict.com) in order to gauge Internet buzz because it allowed me to easily track the number of trailer views for a particular film from week to week. TrailerAddict is a site that offers high-definition trailers through its own custom player. The site is continually updated with new trailers and features various versions of trailers and TV spots for each film. TrailerAddict also features a continually updated list of the

Top 150 most viewed trailers. Next to each film title on the list is the corresponding number of views, an aggregate number of all different trailer versions. It is this number that I noted for each film from week to week.^{6 7}

While the number of trailer views certainly encapsulates awareness within the marketing acronym AIDA (Awareness, Interest, Desire, Action), I argue here that it also encapsulates considerable elements of interest as well. The very name of the site, “Trailer Addict,” suggests an intended audience of those interested in the trailers themselves. In prelease, if a trailer engages the viewer, that viewer is almost certainly going to be interested in the film. The fact that Trailer Addict is a site devoted almost entirely to trailers and not to movie news or information (though there is a small movie newsfeed towards the bottom of the homepage) suggests visitors who are perhaps already familiar with the movies whose trailers they view. Consequently, I argue that the number of trailer views (*TrailerAddict*) is indicative of more than awareness but also of a good trailer that is succeeding in generating interest. I accordingly use it in my analysis as a measure of interest.

3.4.2 Message Board Comments

Another major component of Internet buzz is online discussion and chatter of movie fans. There are numerous blogs and discussion boards where bloggers and posters comment both positively and negatively on movies. The website I ultimately chose to use in my analysis was ComingSoon.net. ComingSoon.net is an all-encompassing movie site that features news, reviews, previews, discussion boards, clips, etc. Like IMDb, ComingSoon.net also has a web page featuring news and clips for each film (In the case of ComingSoon.net, only upcoming and

⁶ Not all of the films in my analysis were included on the Top 150 list. As a default, films not included were assigned the number of views for the 150th film.

⁷ Twice during my analysis, the number of views on the Top 150 were reset. As a result, I divided the aggregate totals by the number of weeks since the reset, and used the average per week number in my analysis.

recently released films have their own web pages.) While IMDb has a full message board on each film's page, ComingSoon.Net has a comments section instead. Though this may inhibit discussion, it provided me with a straightforward way to manually record the number of message board comments. For all of the films in my sample, I simply recorded the number but not the valence of the comments on their respective pages from week to week.

Like the number of trailer views, the number of message board comments not only measures awareness, but it also goes further as a gauge of interest. The action itself of submitting a comment on a film's message board suggests a level of interest beyond just awareness since the commenter is actively responding and putting in the additional effort to express his/her thoughts on the film. Of course, there will be comments among these expressing negative sentiments, but in prerelease these comments are arguably more likely to be fewer in number, unless the film's concept and advertising are particularly egregious. I argue here that, at least in pre-release, if commenters take the time to submit a comment, they are more likely to express excitement than negative sentiment. Accordingly, in my analysis I utilize the number of comments (*ComingSoon*) as a measure of interest and not just of awareness.

3.4.3 Prerelease buzz votes

Another innovation afforded by the Internet for movie fans is the ability to purchase tickets online in advance of the theater at websites like Fandango and MovieTickets.com. On Fandango, each film has its own web page where moviegoers can learn where it is playing in their area and purchase tickets. Also on each individual film's page are trailers, clips, photos, critic reviews, fan reviews, and user ratings. On the user ratings page, there is a section where users can vote "Can't Wait" or "Don't Care" with regards to an upcoming film. Unlike the number of trailer views and the number of message board comments, these prerelease buzz votes

provide an indication of desire and not just awareness or interest. As a result, I divided this buzz measure into two variables. The total number of Fandango buzz votes (*TotalFandango*) will be used to measure awareness and the percent of votes for “Can’t Wait” (*%Can’tWait*) will be used as a measure of desire. Though the prerelease buzz poll requires an active response, the choice of “Can’t Wait” suggests a strong desire while a choice of “Don’t Care” suggests a very firm denial of both interest and desire. As a result, I consider the total votes here to be a pure measure of awareness. Every week, I recorded the number of total votes as well as the number of “Can’t Wait” and “Don’t Care” votes for each film.

Also of note, this variable is the closest to the point of purchase for movie fans, which should hypothetically increase its correlation to opening weekend box office. The only caveat to this variable is that before one places a vote of desire for a particular film, one sees the percentages of how previous users have voted. As a result, new users’ votes may to some extent be influenced by the votes of previous users, thus creating a herding effect.

3.5 Independent Variables: Quality Variables

3.5.1 Rotten Tomatoes Critic Ratings

Rotten Tomatoes is a popular website that features movie news, trailers, interviews and information but is perhaps best known for its aggregate listing of critics’ reviews. The staff at Rotten Tomatoes scours the Internet and gathers reviews for films both from professional reviewers from major media outlets like *The New York Times*, *Variety*, and *Rolling Stone* as well as from reviewers from online film societies. For new and major releases, often over 200 critic reviews are assembled. The Rotten Tomatoes site derives its name from the historical practice of throwing tomatoes at bad acts and in keeping with that theme, the staff classifies each review on

its site (if the critic has not done so himself/herself) as either “Fresh” (Recommended) or “Rotten” (Not Recommended). However, since my purposes are to gauge the perceived quality of the film and not the percentage of recommendation, I decided to track the average critics’ rating. On each film’s respective page on Rotten Tomatoes, an average critics’ rating from 1-10 is listed. Editors at the site convert each applicable review’s original rating, whether it is out of four stars as is Roger Ebert’s custom or a letter grade, to a 1-10 scale and then calculate the average. For my analysis, I recorded this number for each film on its respective opening weekend of release and used it as one of my gauges for film quality. The “worst” film in my sample according to the critics was *Street Fighter: The Legend of Chun-Li* while the “best” film was *Coraline*.

Table 7: Critic Ratings Breakdown

Mean	5.16
Standard Error	0.15
Median	5
Standard Deviation	1.19
Minimum	2.3
Maximum	7.7

3.5.2 Rotten Tomatoes User Ratings

The second gauge I used to assess film quality was the average user rating from Rotten Tomatoes. It probably is not unreasonable to suggest that users and critics are at the very least approaching films from different backgrounds (For one, users pay to see movies while critics are often paid to see movies and write about them) and thus may evaluate film quality on different criteria (Hypothetically, users might evaluate movies on entertainment value while critics might be more inclined to judge a film on its artistic merit). To obtain a complete view of film quality, I decided to track both critic and user ratings alike.

Like other popular movie sites, Rotten Tomatoes has a section where users can vote as a way of expressing their sentiment towards a particular film. Like IMDb, Rotten Tomatoes also uses a scale of 1-10. For each film in my sample, I recorded the average user rating listed after each film’s respective opening weekend. Surprisingly enough, users agreed with the critics on what the “best” and “worst” films in my sample were and also chose *Coraline* and *Street Fighter: The Legend of Chun-Li* respectively.

Table 8: User Ratings Breakdown

Mean	6.49
Standard Error	0.14
Median	6.6
Standard Deviation	1.07
Minimum	3.8
Maximum	8.5

4. Analysis and Results

4.1 Linear Regression: Traditional Variables

In the first part of my analysis, I ran a simple linear regression with domestic opening weekend box office as my dependent variable and budget, star power, MPAA rating, genre, and sequel categorization as my independent variables. The only statistically significant variables in this regression were a film’s budget and categorization as a sequel (Each was significant at a 1% significance level). These results are not particularly surprising and are consistent with some of the prior research. In his study, Ravid (1999) found that budgets and sequels were linked to higher revenues, but not higher returns. Terry et al. (2005) found similar evidence that both budget and sequel categorization were statistically significant predictors of total domestic box

office gross. Simonoff and Sparrow (2000) as well found evidence that sequels outperformed the average film.

Table 9: Linear Regression Results - Predicting Opening Weekend Box Office Using Traditional Variables

<i>Variable</i>	<i>Coefficient</i>	<i>P-value</i>
Intercept	7,889,178	0.489
Rating	-925,932	0.737
Budget	252,437	0.000
Star Power	-27,961	0.913
Sequel	14,935,850	0.009
Action	-2,804,121	0.621
Comedy	1,574,703	0.759
Animated	-4,301,060	0.603
Horror	10,388,148	0.154

Adj. R square= .3665 F-Value= 5.41

The positive relationship between budgets and opening weekend box office makes intuitive sense as large budgets are generally associated with so-called “event” films, action or comic book films with extravagant special effects or children’s films with state of the art computer generated animation intended draw in huge box office gross. To obtain their desired levels of box office, studios typically support their big budget films with hefty advertising expenditures and release them on a vast number of screens across the country. With regards to sequels, since studios generally greenlight their production based on the success of an original film, they know they already have a built-in audience for these particular films, an elusive luxury not afforded to all films. The continued success of sequels, as evidenced once again in my sample, will ensure that their presence in the marketplace will only continue.

It is also worth noting that I did take into the account the number of theaters in which a film was released in its opening weekend, but I did not include it as a dependent variable because of its endogeneity (Studios determine a film’s width of distribution based on their own estimates

of box office gross.) Instead, to account for width of release, I divided opening weekend box office by number of theaters to obtain box office per theater and then reran the regression using this as my dependent variable. This regression yielded similar results as budget and sequel were once again the only statistically significant variables. This makes sense as the correlation between opening weekend box office and opening weekend box office per theater was .93 (Films with wider releases also had higher revenues per theater). Because of this high correlation, opening weekend box office was used as the dependent variable for the remaining regression analyses.

4.2 Buzz Variables

4.2.1. Correlation

For my research, I consulted my selected buzz sources (TrailerAddict, ComingSoon.net, and Fandango) to observe and record the values of their respective buzz variables (number of trailer views, number of message board comments, total votes of desire, and percentage of desire) in each of the three weeks prior to a film's release and on the week of release. To distinguish between data points at different weeks prior to release, I accordingly labeled each as T-3 (3 weeks prior to release), T-2, T-1, and T-0 (week of release). Once I gathered the buzz data for all 62 films in my sample, I measured the correlation between the T-3, T-2, T-1, and T-0 data points across each buzz variable. As the following correlation matrices illustrate, across each buzz variable, the T-3 data point is very highly correlated with the T-data points in following weeks. (Note: For the TrailerAddict variable, the correlation measured accounts for only 56 of the 62 films in my sample. Since the TrailerAddict variable was added later into the research process, the first 6 films in my sample are missing data points for TrailerAddict).

Table 10: Correlation across T-3, T-2, T-1, T-0

Comingsoon.Net

	T-3	T-2	T-1	T-0
T-3	1			
T-2	0.9998	1		
T-1	0.9996	0.9998	1	
T-0	0.9986	0.9990	0.9993	1

TrailerAddict

	T-3	T-2	T-1	T-0
T-3	1			
T-2	0.949	1		
T-1	0.915	0.963	1	
T-0	0.882	0.933	0.984	1

Total Fandango Votes

	T-3	T-2	T-1	T-0
T-3	1			
T-2	0.985	1		
T-1	0.954	0.984	1	
T-0	0.933	0.963	0.990	1

Fandango % "Can't Wait"

	T-3	T-2	T-1	T-0
T-3	1			
T-2	0.975	1		
T-1	0.937	0.987	1	
T-0	0.914	0.974	0.995	1

Particularly of significance are the high correlations between the T-3 and T-0 data points: .9986, .882, .933, and .914 across Comingsoon.net, TrailerAddict, Total Fandango votes, and Fandango % “Can’t Wait” respectively. This is notable because if these buzz variables are indeed significant and do provide predictive value, this predictive value is available at least three weeks prior to release. This three week window before opening weekend is also notable because it is generally in this timeframe that studios escalate their TV advertisements and promotions. Because of its significance and because it is the earliest data point I gathered, the T-3 data point will be used for each buzz variable in the remaining regression analyses. (The only exception will be with regards to the TrailerAddict buzz variable because of the missing data points. In order to include all 62 films in my sample, the T-0 data point will be used. T-0 is .8820 correlated with T-3). In order to attain a sense of scale for each buzz variable used in the following analyses, the following summary statistics are provided.

Table 11: Summary Statistics for each Buzz Variable

	<i>TrailerAddict</i>	<i>ComingSoon</i>	<i>TotalFandango</i>	<i>%Can'tWait</i>
Unit of Measure	# of trailer views	# of comments	# of votes	percent
Mean	5933.81	78.21	522.34	48.24%
Standard Error	974.68	15.82	49.62	1.99%
Median	3480.00	36.50	430.50	48.50%
Standard Deviation	7674.61	124.55	390.67	15.66%
Minimum	568.00	2.00	35.00	15.00%
Maximum	45865.69	594.00	1778.00	79.00%

Before assessing the predictive value of the buzz variables, I also measured the correlation between each buzz variable.

Table 12: Correlation across Buzz Variables

	<i>TotalFandango</i>	<i>%Can'tWait</i>	<i>TrailerAddict</i>	<i>ComingSoon</i>
TotalFandango	1			
%Can'tWait	0.41	1		
TrailerAddict	0.46	0.42	1	
ComingSoon	0.50	0.43	0.52	1

When paired up, the buzz variables do indicate some correlation with one another. The highest correlated pair was *TrailerAddict* and *ComingSoon* (.519) and the lowest correlated pair was *TotalFandango* and *%Can'tWait* (.412). However, since this matrix only provides information in separate pairs, I decided to run four different regressions, each one using one of the buzz variables as the dependent and the remaining buzz variables as the independents. By evaluating the explanatory power of the buzz variables on each other, I can assess whether any of them was considerably capturing the effects of the others. The R-squares of each regression were as follows:

Table 13: Linear Regression Results - Explaining each Buzz Variable with the Remaining Buzz Variables

Dependent Variable	Independent Variables	P-Values	R Square	F- Value
TrailerAddict	ComingSoon TotalFandango %Can'tWait	0.045 0.052 0.124	0.327	9.38
ComingSoon	TrailerAddict TotalFandango %Can'tWait	0.045 0.038 0.054	0.357	10.73
TotalFandango	TrailerAddict ComingSoon %Can'tWait	0.052 0.038 0.136	0.328	9.46
%Can'tWait	TrailerAddict ComingSoon TotalFandango	0.124 0.054 0.136	0.292	7.96

These results indicate that none of the buzz variables is significantly capturing the effects of the rest, thus illustrating that they are perhaps providing unique information, such as awareness, interest, and desire.

4.2.2 Linear Regression

To begin analyzing the value of the buzz variables, I first ran four different regressions, each incorporating a different buzz variable alongside the traditional independent variables of genre, budget, star power, MPAA rating and sequel. While budget and sequel remain statistically significant in each regression, each buzz variable also proves statistically significant and is associated with higher opening weekend box office (*TrailerAddict*, *ComingSoon* and *%Can'tWait* are all significant within a 1% significance level. *TotalFandango* is significant within a 5% level.) These results indicate that, at least separately, these buzz variables as proxies

for prerelease awareness, interest, and desire do provide considerable predictive value in addition to that provided by the individual characteristics of the films.

Table 14: Linear Regression Results – The Additional Predictive Value of each Individual Buzz Variable

<i>Variable</i>	<i>Coefficient</i>	<i>P-value</i>	<i>R Square of Regression</i>	<i>F-Value</i>
TrailerAddict	1,119	0.00	0.592	10.84
ComingSoon	70,575	0.00	0.579	10.33
TotalFandango	11,168	0.03	0.437	6.27
%Can'tWait	56,705,335	0.00	0.550	9.29

I then followed up these individual regressions with a simple linear regression that incorporated each of the buzz variables at the same time. The results from this regression reveal that *TrailerAddict* is significant within 1%, *ComingSoon* and *%Can'tWait* are each significant within 5% and that *TotalFandango* is not statistically significant. Budget and Sequel also remain statistically significant and, somewhat counter intuitively, the Action genre becomes statistically significant at a 1% level with a negative coefficient. (This rather surprising result regarding the Action dummy variable may stem from my limited sample. If I was able to gather data for an entire year, thereby including the summer blockbuster season which is typically filled with hugely successful action films, this unexpected result likely may not have appeared.)

Table 15: Linear Regression Results – Predicting Opening Weekend Box Office with both Traditional and Buzz Variables

<i>Variable</i>	<i>Coefficients</i>	<i>P-value</i>
Intercept	178,109	0.984
Rating	-2,411,803	0.232
Budget	136,156	0.005
Star Power	-35,996	0.852
Sequel	14,061,187	0.003
Action	-12,152,723	0.006
Comedy	1,070,088	0.774
Animated	-2,461,055	0.675
Horror	1,636,656	0.764
TrailerAddict	696	0.004
ComingSoon	33,957	0.035
TotalFandango	-617	0.887
%Can'tWait	31,702,196	0.012

Adj. R Square = .6790 F-Value = 11.75

With regards to the buzz variables, these results are interesting on a couple of levels. First of all, by incorporating the buzz variables along with the traditional variables, I was able to add a great deal of explanatory power to the model, increasing the R^2 dramatically from .3665 to .6790. These results illustrate that the buzz data does indeed provide considerable predictive value and that this value is available at least three weeks prior to opening weekend (and perhaps even earlier but that is a matter for another study). They also suggest that studios can find easily accessible, valuable and perhaps even actionable information by tracking the level of internet buzz for their films in the weeks prior to their release. By obtaining information not only on awareness but on consumer interest and desire, studios can obtain a useful signal on the prospects of their upcoming films. Even further, by learning how consumers are responding to advertisements, in addition to their level of awareness, studios can use this information on buzz in aid of decisions not only to increase or decrease advertising expenditures but also in deciding whether or not to change the content of the ads themselves. Overall, this prerelease buzz data can

provide studios with a gauge of consumer sentiment towards their films before they have actually been seen, especially useful information when considering the importance of opening weekend box office.

In section 3, I categorized each buzz variable as a different point in the marketing acronym, AIDA (Awareness, Interest, Desire, Action). These results may suggest that awareness (*TotalFandango*) alone is not significant for box office success and that awareness needs to be coupled with elements of interest (*TrailerAddict* and *ComingSoon*) and desire (*%Can'tWait*) in order to generate substantial box office. This result is important because it potentially suggests that increasing advertising expenditures for the sake of increasing awareness may not by itself be effective in getting moviegoers into the seats and that these expenditures would perhaps be better served in support of a creative and engaging advertising campaign. Simply getting the message out there may not suffice; the message has to be one audiences are interested in hearing.

However, since advertising is unaccounted for in this regression, I do not mean to suggest a link between advertising and buzz, though that link may seem intuitive. Instead, I mean simply to offer a possible and perhaps viable interpretation to the results. In section 4.4, I attempt to evaluate whether the individual characteristics of a film (budget, sequel, star power, etc.) are associated with buzz.

4.3 Film Quality: User and Critic Ratings

Utilizing the ratings data I collected from Rotten Tomatoes, I then conducted a series of regressions to assess whether film quality, as expressed as word of mouth by fans and critics, is relevant to box office success. Before I ran the regressions, however, I determined the correlation between the user and critic ratings to be .66. While notable, this correlation may be somewhat inflated by the structure of the Rotten Tomatoes' site. When a visitor navigates to a film's page

on the Rotten Tomatoes' site, the default page features the critic ratings and critic percentage on the Tomatometer. While users can provide their rating on this page, users may perhaps anchor their ratings by the critics' rating.

In my first regression, I incorporated the ratings data into the traditional model and my results showed that neither critic nor user ratings were significant in relation to box office and that they provided no additional predictive value. In my second regression with ratings, I incorporated each into my buzz model and, once again, neither proved to be statistically significant.

Table 16: Results from Film Quality Regressions

Trad. + RT Critics and RT Users

<i>Variable</i>	<i>Coefficients</i>	<i>P-value</i>
Intercept	-1,230,341	0.936
Rating	-1,319,739	0.679
Budget	197,895	0.002
Star Power	51,410	0.845
Sequel	18,275,542	0.003
Action	1,056,178	0.866
Comedy	2,176,021	0.691
Animated	-668,908	0.938
Horror	11,906,644	0.136
RTCritic	2,289,544	0.349
RTUser	-257,701	0.918

Adj. R Square = .3189

F-Value = 3.86

Trad. + RT Critics + RT Users + Buzz

<i>Variable</i>	<i>Coefficients</i>	<i>P-value</i>
Intercept	-13,754,937	0.223
Rating	-4,217,663	0.074
Budget	92,138	0.063
Star Power	20,258	0.919
Sequel	14,638,065	0.003
Action	-8,799,891	0.061
Comedy	2,547,215	0.517
Animated	-2,053,450	0.734
Horror	4,776,432	0.415
RTCritic	2,832,922	0.120
RTUser	549,766	0.764
TrailerAddict	737	0.004
ComingSoon	41,058	0.017
TotalFandango	-3,183	0.490
%Can'tWait	34,589,324	0.008

Adj. R Square = .6709

F-Value = 9.88

In terms of critic ratings, these ratings are consistent with some of the prior research. Eliashberg and Shugan (1997) looked at the percentage of positive and negative critic reviews (I analyzed average critic rating) and found that both were indeed statistically significant predictors

of overall box office. However, they did not find evidence that these measures of valence were statistically significant towards early week box office (weeks 1-4), which corresponds with the results of my opening weekend analysis. Ravid (1999) found no significant positive relationship at all between percentage of positive reviews and total revenues. On the other hand, using Siskel and Ebert's "Thumbs Up" as their measure of critic reviews, Reinstein and Synder (2005) did find evidence, albeit admittedly marginal, that critic reviews were statistically significant influencers of opening weekend box office. When they broke down their analysis by category, they found stronger evidence to indicate critics were statistically significant influencers of narrowly-released films and dramas. Basuroy et al. (2003) found similar evidence that both positive and negative critic reviews are significantly correlated with box office, including in the early weeks. Even further, they found that in earlier weeks negative reviews hurt box office revenue more than positive reviews helped. Lastly, they found that both star power and budget were significant in lessening the impact of negative reviews.

Since my analysis only uses opening weekend box office as the dependent variable, it is not especially surprising that neither quality variable contributed significant explanatory power. First of all, it seems relatively fair to say that critics do not share the same taste as the rest of the movie-going populace. Despite critics' glowing recommendations or stern early warnings, there is some evidence both from the prior literature and my analysis to suggest that moviegoers see what they want to see, irrespective of critics' opinion. As evidenced by Reinstein and Synder (2005), the instances where critics do have more influence is with regards to the more serious, drama films, where perceived merit and quality are more likely to be a criterion in the moviegoers' decision process. Event films and blockbusters, on the other hand, are much more likely to be immune to any critical vitriol as indicated by Basuroy et al. (2003). The textbook

example supplied by my sample is that of *Fast & Furious*, which attained the highest opening weekend box office among the 62 films I analyzed. A critic rating of 4.4 (not to mention a 24% recommendation percentage, though this of course was not included in the primary analysis) was not able to get in the way of its \$70.95 million opening. On other hand, the highest critically rated film in my sample, *Coraline*, with a critic rating of 7.7 only had an opening of \$16.85 million.

In the case of the users' ratings, it does not come as much of a surprise that these ratings were not significant towards opening weekend box office since, after all, users cannot decide whether they like a movie or not until they have seen it. For example, this would intuitively explain the number one openings of over \$30 million for films like *Four Christmases* and *The Day the Earth Stood Still*, who each had relatively low user ratings of 5.3 and 4.9 respectively.

If anything, coupling the ratings data with the buzz data only further highlights the critical importance of creating an image of quality prior to initial release for experiential products like movies. For example, after months and even years of production and after millions of dollars invested, if a studio realizes they might have a "bad" film on their hands, one that it strongly feels will not strike a chord with audiences, it can try its best to earn as much return as it can on opening weekend before negative post-release word of mouth spreads. As this analysis has illustrated, prerelease buzz is indeed positively associated with opening weekend box office. Accordingly, in this case the studio should try to generate buzz at a low cost as best it can so that it can successfully "dump and run" on opening weekend. A similar buzz strategy would also apply to studios who feel like they have a "good" film or box office hit on their hands. Because of the significant portion of total box office revenues that opening weekend constitutes, a "good" film with little prerelease buzz risks missing out on considerable revenues. Even further, it may

also lose out on the even stronger word of mouth that could potentially come with having a successful opening weekend, simply as result of more people having seen the film. Any good word of mouth that is generated in a subpar opening weekend can only go so far as more and more competing films are released each week. Since pre-release buzz has been shown to be significant towards opening weekend box office, the clear question that then emerges is “What contributes to buzz in the first place?” In the next section, I attempt to provide an initial answer to this question.

4.4 What Contributes to Buzz?

In order to evaluate what film characteristics contribute to buzz, I ran four simple linear regressions, each using one of the buzz variables (*TrailerAddict*, *ComingSoon*, *TotalFandango*, *%Can'tWait*) as a dependent variable and the traditional variables (budget, genre, star power, MPAA rating, sequel) as the independents. The results were as follows:

Table 17: Linear Regression Results - Predicting Buzz with the Traditional Variables

<i>Dependent Variable</i>	<i>Significant Predictor(s)</i>	<i>P-Values</i>	<i>R-Square</i>
Trailer Views	Budget	0.004	0.152
Message Board Comments	Sequel	0.016	0.24
	Action	0.042	
Total Buzz Votes	None	N/A	0.155
% "Can't Wait"	Budget	0.004	0.366
	Action	0.008	

Budget, one of the two traditional variables that was a significant predictor of opening weekend box office in the previous regressions, is unsurprisingly a significant predictor of *TrailerAddict* and *%Can'tWait*, measures of both interest and desire, as well. Intuitively, this makes sense as higher budgets, generally associated with ‘event’ films, not only contribute to opening weekend box office, but they also contribute to the buzz that does so as well. Likewise, a similar interpretation seems to apply to sequel, the other significant box office predictor. The

categorization as a sequel contributes to opening weekend box office and is also a significant predictor of the buzz variable, *ComingSoon*, a contributor itself. Since sequels already have built-in audiences, it is easy to see why sequels would be associated with higher levels of buzz and, specifically, interest. These results concerning budget and sequels may reinforce the importance for studios and advertisers of crafting the image of an ‘event’ film in their campaigns.

Seemingly contradictory to the results in Table 14, in which the Action dummy variable was significant and negatively associated with opening weekend box office, are the results here that reveal action films to be statistically significant and positively associated with *ComingSoon* and *%Can'tWait*, positive predictors of opening weekend box office. However, there were a number of action films in my sample that performed relatively poorly, yet had considerable buzz in prerelease, especially on *Comingsoon.net* and *Fandango.com*. For example, a film like *Punisher: War Zone* had high buzz values of 419 comments on *ComingSoon* (Mean = 78.2) and 62% on *%Can'tWait* (Mean = 48.2%), but only grossed \$4.3 million on opening weekend.

5. Managerial Implications

As illustrated in this study, incorporating internet buzz provides considerable explanatory power when predicting opening weekend box office. Given these results, studios, if they are not already doing so, should monitor and remain very conscious of the buzz and chatter that takes place online with regards to their films. Whether it is in a similar manner to what I have done in this study or in a more scientific fashion, perhaps through special computer programs, tracking internet buzz data could potentially be used alongside the tracking reports received by studios from firms like Nielsen NRG and MarketCast. Like the data provided by these third party sources, buzz data also can be broken down into categories of awareness, interest and desire.

While demographic data may be more difficult to obtain for buzz data, the additional value of monitoring buzz data could be in its ability to provide even earlier predictions.

Movies typically appear on these tracking surveys three weeks prior to release, the same three week window I tracked for the buzz data in my study. These tracking surveys provide studios with estimates of how well their advertising is raising awareness, generating interest, and reaching certain demographics. This data allows studios to obtain a highly accurate estimate of opening weekend box office by the day of release.⁸ However, if studios had more accurate estimates further in advance of release, they would have more time to take action, either through increasing/decreasing ad expenditures or by altering the message of their campaign. While I did not track buzz data earlier than three weeks in advance, the very high correlations between my T-3 and T-0 variables suggest that this predictive value may be available well before this three week window. As I mention in the next section, further research that tracks buzz data even further in advance of release could prove very valuable for studios.

Another implication of this study was the suggestion that awareness alone may not be sufficient in generating box office. Ultimately, my results illustrated that only those buzz variables associated with interest and desire were significant and not the one associated with awareness. For studios, this may further emphasize the importance of crafting advertising messages and campaigns that engage consumers and generate interest and desire. Simply increasing ad spending so that more consumers become aware of your film may not alone get the job done. Studios need to be diligent in monitoring whether or not their advertising efforts are striking a chord with their desired audience.

⁸ Horn, John. "The Inside Track." Newsweek.com. 7 Feb. 2002. <www.newsweek.com/id/63415>.

6. Limitations and Further Research

To add onto this study, one area of further research would be to use alternative websites and thus alternative variables as measures of Internet buzz. Additional research that utilizes alternative gauges of buzz and takes into account movie websites not included in this study can help assess the extent to which the results found here are applicable to movie buzz in general. Utilizing the variables within this analysis, one can also add another measure of desire or preference by examining the comments on ComingSoon.net and classifying each as positive or negative in a similar fashion to Liu's study with Yahoo message board comments (2006).

Given my limited timeframe, I was only able to track data for a sample over five months. For more comprehensive results, it would also be beneficial to examine a larger sample, particularly one that includes movies during the summer blockbuster season. Since these are the type of films likely to have considerable buzz, it would be interesting to see how including these films affects the results.

Another avenue of additional research that would be of great interest would be to examine buzz even further back in time. In my study, I only tracked buzz three weeks prior to a film's release. As my analysis revealed, the data from three weeks prior to opening weekend was highly correlated with the following weeks. It would be very interesting to see just how far back this high correlation extends and to determine how far in advance of opening weekend does buzz begin to build. If it turns out that buzz data does provide predictive value in additional weeks prior to release, studios could potentially obtain actionable information with even more time to utilize it.

Finally, my study here offered a preliminary attempt to evaluate what film characteristics may contribute to buzz. Much more research can be done in this regard, particularly a more

focused look on what actions a studio itself can take to generate buzz. I myself was not able to obtain accurate data regarding advertising expenditures, but further research on advertising and its relationship to buzz would be especially beneficial for studios.

7. Conclusion

Not only has the Internet facilitated the spread of word of mouth and buzz in the movie industry, it has also provided another area for both studio and researchers to assess consumer sentiment. For the movie industry, since their product is experiential, it is of even more importance to gauge word of mouth and buzz prior to a movie's release. In this study, I attempted to capture prerelease Internet buzz in the movie industry and to evaluate its relationship with opening weekend box office. Using trailer views, message board comments, and prerelease votes of desire to generate four different independent variables, each categorized as either awareness, interest, or desire, I ran a linear regression to evaluate the predictive value of Internet buzz. In my analysis, I find that buzz, along with a film's budget and its categorization as a sequel, is indeed statistically significant in predicting opening weekend box office. When I classify the buzz variables as proxies for awareness, interest and desire, I find that only those associated with interest and desire are statistically significant. On the other hand, awareness alone is not, thus perhaps highlighting the importance of crafting engaging advertising campaigns over simply increasing awareness.

In the second part of my analysis, I evaluate the relationship of critic ratings and user ratings to opening weekend box office. Consistent with the prior research of Eliashberg and Shugan (1997) and Ravid (1999) but contrary to that of Basuroy et al. (2003) and Reinstein and Snyder (2005), I find no significant relationship between critic ratings and opening weekend box office. Rather unsurprisingly, I also find that user ratings, a post-release measure, are not

significant towards opening weekend box office. These results potentially reinforce the importance of generating prerelease buzz. In the final part of my analysis, I also analyze what characteristics of a film may be instrumental in generating buzz. My findings show that higher budgets, sequels and action films all generate higher levels of buzz.

Bibliography

- Basuroy, Suman, Subimal Chatterjee and S. Abraham Ravid (2003), "How Critical Are Critical Reviews? The Box Office Effects of Film Critics, Star Power and Budgets?" *Journal of Marketing*, Vol. 67. 103-117. Oct. 2003.
- Box Office Mojo. <www.boxofficemojo.com>.
- ComingSoon.net. <www.comingsoon.net>.
- Dellarocas, Chrysanthos, Xiaoquan (Michael) Zhang, and Neveen F. Awad (2007). "Exploring the Value of Online Product Reviews in Forecasting Sales: The Case of Motion Pictures." *Journal of Interactive Marketing*, Vol. 21, No. 4, 23-45. Nov. 2007.
- Duan, Wenjing, Bin Gu, and Andrew B. Whinston (2005). "Do Online Reviews Matter? – An Empirical Investigation of Panel Data." *Decision Support Systems*, Vol. 45, No. 4, 1007-1016. Nov. 2008
- Eliashberg, Jehoshua and Steven Shugan (1997). "Film Critics: Influencers or Predictors?" *Journal of Marketing*, Vol. 61, No.2, 68-78. Apr. 1997.
- Fandango. <www.fandango.com>.
- Godes, David and Dina Mayzlin (2004). "Using Online Conversations to Study Word-of-Mouth Communication." *Marketing Science*, Vol. 23, No. 4, 545-560. Fall 2004.
- Hayes, Dade and Jonathan Bing. *Open Wide: How Hollywood Box Office Became A National Obsession*. New York: Hyperion. 2004.
- Horn, John. "The Inside Track." *Newsweek.com*. 7 Feb. 2002.
<www.newsweek.com/id/63415>.
- The Internet Movie Database. <www.imdb.com>.

- Lavidge, Robert J. and Gary A. Steiner (1961). "A Model for Predictive Measurements of Advertising Effectiveness." *The Journal of Marketing*, Vol. 25, No. 6, 59-62. Oct. 1961.
- Litman, Barry R. (1983). "Predicting Success of Theatrical Movies: An Empirical Study." *Journal of Popular Culture*, Vol. 16, No. 4. 159-175. Spring 1983.
- Liu, Yong (2006). "Word of Mouth for Movies: Its Dynamic and Impact on Box Office Revenues." *Journal of Marketing*, Vol. 70, No. 3, 74-89. Jul. 2006.
- McNary, Dave and Pamela McClintock. "High hoped for 'Madagascar' sequel." *Variety*. Nov. 6, 2008. <<http://www.variety.com/article/VR1117995403.html?categoryid=13&cs=>>
- Ravid, S. Abraham (1999). "Information, Blockbusters, and Stars: A Study of the Film Industry." *Journal of Business*. Vol. 72, No. 4, 463-492.
- Reinstein, David A. and Christopher M. Snyder (2005), "The Influence of Expert Reviews on Consumer Demand for Experience Products: A Case Study of Movie Critics." *Journal of Industrial Economics*, Vol. 53, No. 1, 27-51. Mar. 2005.
- Simonoff, Jeffrey S. and Ilana R. Sparrow (2000). "Predicting Movie Grosses: Winners and Losers, Blockbusters and Sleepers." *Chance*. Vol. 13, No. 3, 15-24.
- "Star Currency: Complete Methodology." *Forbes.com*. Feb.10, 2009.
<http://www.forbes.com/2009/02/06/forbes-star-currency-methodology-business-media-star-currency-09_0210_methodology.html>
- Strong, E.K. (1925). "Theories of selling." *Journal of Applied Psychology*. Vol. 9, No. 1, 75-86. Mar. 1925.
- Terry, N. and Butler M. and De'Armond D. (2005). "The Determinants of Domestic Box Office Performance in the Motion Picture Industry." *Southwestern Economic Review*, Vol. 32, No. 1, 137-148.

Trailer Addict. <www.traileraddict.com>.

Appendix

Actual vs. Predicted Opening Weekend Box Office (Buzz Model)

<i>Week</i>	<i>Movie</i>	<i>Actual</i>	<i>Predicted</i>
1	Role Models	\$19,167,085	\$15,956,608.96
1	Madagascar 2	\$63,106,589	\$56,954,802.32
1	Soul Men	\$5,401,605	\$8,825,820.16
2	Quantum of Solace	\$67,528,882	\$58,056,654.07
3	Bolt	\$26,223,128	\$32,953,533.75
3	Twilight	\$69,637,740	\$62,098,390.57
4	Australia	\$14,800,723	\$28,635,569.82
4	Four Christmases	\$31,069,826	\$25,255,725.51
4	Transporter 3	\$12,063,452	\$23,592,663.61
5	Punisher: War Zone	\$4,271,451	\$18,857,326.98
5	Cadillac Records	\$3,445,559	\$21,037.71
6	Delgo	\$511,920	\$3,827,674.95
6	The Day the Earth Stood Still	\$30,480,153	\$26,534,820.40
6	Nothing like the Holidays	\$3,531,664	\$9,315,389.97
7	Seven Pounds	\$14,851,136	\$28,689,132.24
7	Tale of Despereaux	\$10,103,675	\$22,988,664.68
7	Yes Man	\$18,262,471	\$26,112,934.44
8	Marley and Me	\$36,357,586	\$20,058,287.59
8	Bedtime Stories	\$27,450,296	\$26,432,072.54
8	The Curious Case of Benjamin Button	\$26,853,816	\$39,722,778.85
8	Valkyrie	\$21,027,007	\$22,711,940.57
8	The Spirit	\$6,463,278	\$9,939,624.85
8	Doubt	\$5,339,742	\$4,065,968.78
10	Gran Torino	\$29,484,388	\$11,368,495.27
10	Bride Wars	\$21,058,173	\$19,272,404.68
10	The Unborn	\$19,810,585	\$23,413,850.24
10	Not Easily Broken	\$5,314,278	-\$1,215,644.47
11	Defiance	\$8,911,827	\$11,217,207.59
11	Hotel for Dogs	\$17,012,212	\$12,720,461.05
11	My Bloody Valentine 3D	\$21,241,456	\$8,306,700.45

Actual vs. Predicted Opening Weekend Box Office (Buzz Model) (Continued)

<i>Week</i>	<i>Movie</i>	<i>Actual</i>	<i>Predicted</i>
11	Notorious	\$20,497,596	\$9,982,726.36
11	Paul Blart: Mall Cop	\$31,832,636	\$23,753,517.54
11	Last Chance Harvey	\$4,299,805	\$6,589,392.10
12	Underworld: Rise of the Lycans	\$20,828,511	\$30,891,196.98
12	Inkheart	\$7,601,379	\$11,687,476.24
12	Revolutionary Road	\$5,185,146	\$10,207,865.75
13	Taken	\$24,717,037	\$3,703,695.07
13	The Uninvited	\$10,325,824	\$11,303,792.94
13	New in Town	\$6,741,530	\$13,341,931.79
14	He's Just Not That Into You	\$27,785,487	\$16,583,737.39
14	Coraline	\$16,849,640	\$13,498,835.46
14	Push	\$10,079,109	\$7,695,852.16
14	Pink Panther 2	\$11,588,150	\$31,637,935.26
15	Confessions of a Shopaholic	\$15,066,360	\$17,064,170.28
15	The International	\$9,331,739	\$6,553,859.71
15	Friday the 13th	\$40,570,365	\$45,699,581.10
16	Fired Up	\$5,483,778	\$5,999,259.73
16	Tyler Perry's Madea Goes to Jail	\$41,030,947	\$25,467,021.20
17	Jonas Brothers: The 3D Experience	\$12,510,374	\$15,491,330.10
17	Street Fighter: The Legend of Chun-Li	\$4,721,110	\$15,514,123.16
18	Watchmen	\$55,214,334	\$59,394,538.37
19	Last House on the Left	\$14,118,685	\$12,771,303.80
19	Miss March	\$2,409,156	\$11,539,556.69
19	Race to Witch Mountain	\$24,402,214	\$14,741,471.14
20	Duplicity	\$13,965,110	\$13,170,974.52
20	I Love You, Man	\$17,810,270	\$19,207,938.40
20	Knowing	\$24,604,751	\$23,730,667.63
21	Monsters vs. Aliens	\$59,321,095	\$45,892,535.84
21	12 Rounds	\$5,329,240	\$8,673,665.52
21	The Haunting in Connecticut	\$23,004,765	\$27,576,451.47