Social TV:
Real-Time Social Media Response to TV Advertising

Shawndra Hill
University of Pennsylvania
The Wharton School
3730 Walnut Street
Philadelphia, PA 19104
shawndra@wharton.upenn.edu

Aman Nalavade
Citadel Investment Group
601 Lexington Avenue
New York, New York 10022
aman.nalavade@gmail.com

Adrian Benton
University of Pennsylvania
The Wharton School of Business
3730 Walnut Street
Philadelphia, PA 19104
adrianb@mail.med.upenn.edu

ABSTRACT
We link the *content* of real-time social media response to characteristics of the Super Bowl 2012 TV advertisements, including their advertisers' social media strategies and the categories of the products being advertised. We analyze millions of social media posts about approximately forty-five advertisements during the 2012 Super Bowl. While there are many studies that focus on the popularity of and sentiment of response to Super Bowl ads, we believe ours is the first to go deeper and analyze the textual "content" to understand better the factors contributing to specific types of social media responses among a number of TV advertising dimensions. In this study, we show that the level of online consumer engagement, measured by attracting new followers on Twitter during the Superbowl, can be linked to whether the brand that advertised had a social media strategy or not. In addition, we show that sentiment response, a measure often use to quantify the effectiveness of TV ads, varies by demographic category.

Categories and Subject Descriptors
H.3.4 [Social Networking]

General Terms
Measurement, Human Factors

Keywords
social media, social TV, television, advertising

1. INTRODUCTION
"Social TV" is the term used to describe the current integration of social media interaction with television programming. Social television has sought to recapture the early days of television, where families gathered in their homes to share the experience of watching television together [1]. With the proliferation of social media applications and smart phone technology, social interaction around television programming can now be shared among millions of viewers all at once. It is estimated that there are an average of ten million public online comments made each day worldwide related to television content [2]. Twitter and other social media platforms have "become an integral outlet for TV viewers who are looking to express themselves while watching broadcasts of their favorite television programs" [3]. This "backchannel" of communication during TV shows has also led to a resurgence in people's interest in watching live shows [3].

For researchers and firms alike, the prevalent question is how to make sense of and derive value from it all. For the first time, advertisers are able to get *real-time feedback* in the form of sentiments (positive, negative and neutral) from large audiences about their products and ads. In addition, networks are able to capture comments from viewers throughout a television show and therefore can ask important questions, such as these: How can programs and advertisers induce engagement from viewers? If an announcer tells a consumer to go to a certain website, Tweet, or purchase a particular item, does that influence the consumer to actually comply? Audience response is both immediate and measurable. The amount of user-generated content (UGC) in the context of TV is immense. The 2012 Super Bowl resulted in 13.7 million tweets alone [4].

While the Super Bowl generated a lot of "buzz", which means a lot of user-generated content about the Super Bowl, the challenge for researchers and industry alike is to be able to move beyond merely tracing the amount of "buzz" generated in order to provide precise content analysis of what is being said by potential consumers in response to what is being shown on TV. In other words, we can begin to understand *"what the buzz is all about."* This will allow for marrying social-TV data about the shows with data about specific brands [1]. The problem is that the raw data is a noisy stream of consciousness. Therefore, sophisticated tools are needed to extract meaningful text-based features before tweets can be used in decision making. This work demonstrates that content analysis can reveal important features about the buzz that can be linked to business outcomes.

Our work falls at the intersection of Information Systems, Computer Science, and Marketing. These fields intersect in a significant amount of recent research aimed at mining publicly available text and network data to predict business outcomes. For example, firms can derive better hotel recommendations from online hotel reviews [5] and social networks can be linked to the spread of information online and by word-of-mouth [6-8] as well as used to predict product adoption [9]. Prior work has also shown that features from social media, in general, and Myspace Music, in particular, can be a predictor of ranked sales [10, 11]. In addition, research has shown a link between drugs and their side-effects from medical discussion board data [12, 13].

Despite the abundance of UGC, and its many uses for business, researchers are only beginning to understand how word of mouth differs across channels [14]. In our work, we try to understand

ADKDD'12, August 12, 2012, Beijing, China.

the link between what is presented on TV and word of mouth online. While our work will build on the recent endeavors to extract value from social media and UGC, it also relates to more traditional advertising work. We will build on the TV advertising [15-21] literature, particularly work that focuses on Super Bowl TV advertising. Prior studies have tried to link characteristics of Super Bowl ads, including how long they were shown and the number of times, to popularity and likeability using focus groups and post-hoc surveys [17, 22]. While we have not found any published academic work linking TV advertising to social media buzz, there has been recent work that tries to link what a particular talk show host says about companies to stock price movement [23].

Our contribution to the current research on Social TV is developing methods for extracting features from the contents of online consumer discussions to predict advertising outcomes. Table 1 illustrates that, while there may be many tweets about a particular brand, they are not all created equal. In fact, we find in our work that very few social media posts about TV advertisements are about details of the products being advertised. Our work will extract the factors of tweets and link the meaningful attributes of the TV advertisements to them.

In this work, we pay particular attention to the content of the social media messages and develop tools to answer the following questions: Which ad characteristics seem to trigger the most tweeting and the longest tweeting half-life? Are the tweets about the product or the creative aspects of ads, and when? Who is tweeting? Is there a demographic bias to the tweeting? Who benefitted the most from the interaction between what is said on TV and Social Media? How did the follower network grow over time? Does including social media content in TV advertising result in more viewer engagement?

With these questions in mind, we show two results: 1) Including social media content in the ad and/or having a social media strategy resulted in more buzz, increased followers and a higher response from female user-generated content providers; and 2) Sentiment response (positive, negative, neutral) varied by demographic category.

Table 1. Examples of tweets about the *product* (left) and about the creative aspects of *ads* (right)

Why are Doritos the best chips in the world?	Doritos commercial was cute. Liked the dog bribing the human.
Just saw my new Honda on the TV for the first time.	Have you seen the Honda ad yet? Total destruction of a Ferris classic.
Buying a Chevy at cars.com.	That cars.com commercial was hilarious.

2. Literature Review

The Super Bowl is a wildly popular television event. The 2012 game drew a record-breaking 111.3 million viewers, making it the most watched television show in US history [4]. In what has now become a familiar tradition, one of the most anticipated aspects of the Super Bowl is not the game itself, but the commercials airing during the game. Advertisers paid a huge premium to capture the attention of the viewers of the 2012 game, paying on average $3.5 million for a thirty-second commercial [24]. Advertisers pay a premium because they believe it is an efficient way to reach millions of both male and female viewers in one stop.

Advertisers also have efficacy objectives. As a result, several companies have emerged to measure ad effectiveness. Companies such as Bluefin Labs, Networked Insights, Social Guide and Trendrr specialize in analyzing these data and generating metrics for clients interested in their "online image.".

One of the earliest attempts to measure the real-time effectiveness of Super Bowl commercials was pioneered by USA Today in 1989 using focus groups in what they coined “the Super Bowl Ad Meter panels” [25]. Today, with the widespread proliferation of smartphones, laptops and tablet computers, combined with the meteoric rise of social media platforms such as Twitter and Facebook, real-time effectiveness of Super Bowl ads can now be measured instantaneously on a grand scale. Companies and advertisers are able to see second-by-second trending of buzz generated by their products. The 2012 Super Bowl peaked at a record-breaking 12,233 tweets per second [26].

The popularity of evaluating advertising campaigns during the Super Bowl is evidenced by academic researchers at both the Wharton School of Business and the Kellogg School of Management attempted to evaluate 2012 Super Bowl ad effectiveness. Kellogg used a panel of students and faculty, much like the Ad Meter panels, measuring the ad effectiveness along a six metric strategic academic framework [27]. Wharton used a “selective crowdsourcing” of industry insiders and academics connected via Twitter, measuring and rating advertisements on simple criteria. In addition, these experts generated additional Twitter comments about the ads [28]. An expert ranking of ads resulted from both the Wharton and Kellogg studies based on the perceived quality of the ads.

Most online rankings were primarily based on popularity measure by overall buzz and overall sentiment. Despite a wealth of popularity-based rankings, little was reported about the content of the buzz. Being able to gauge and understand the contents of the buzz around a product is valuable tool for advertisers to engage consumers better and provide target marketing strategies, maximizing return on investments. This paper attempts to link attributes of the advertisements to the content of the buzz using ad attributes that have been well established in the literature.

2.1 Methods of Evaluating Advertisements

Prior research has yielded several studies measuring the effectiveness of television advertisements. In a study titled, “The Effect of Antismoking Advertisement Executional Characteristics on Youth Comprehension, Appraisal, Recall, and Engagement” [29], television advertisements were measured for comprehension, appraisal, recall, and engagement by varying two executional characteristics: 1) personal testimonial and 2) negative visceral image. The responses to the television commercials were also compared within two audience segments: 1) youth and 2) general audience. The commercials with personal testimonials used emotional appeals such as fear, sadness, or empathy, which were believed to strengthen the relevance and credibility of the commercial. The use of negative visceral images stemmed from previous research on fear appeals [18] and from the hypothesis that this type of intense imagery can reinforce message relevance, credibility, and recall.

In “Recall and Persuasion: Does Creative Advertising Matter?” [21], the authors measured the effectiveness of creative advertising with regard to its ability to affect recall, brand attitude, and purchase intent. The authors referred to three prior studies to determine whether an ad is “creative.” One study by Kover, Goldberg, and James [30], used participants’ measure of creativity in terms of describing an ad as old/new or dull/exciting. In another study by Ang and Low [31], the authors classified the advertisements themselves as creative in terms of novelty, meaningfulness, and emotional content of user-generated response to TV. The third study, by Stone, Besser, and Lewis [32], linked ad creativity and likeability, finding that seventy percent of liked commercials were considered creative and only forty-six percent of disliked commercials were judged to be creative. The results showed that creative commercials help facilitate unaided recall and the effect persists over a one-week time period. However, creativity does not change a respondent's recall ability using an aided recall process. Furthermore, creativity did not seem to have an impact on purchase intent or brand attitude.

Another study, "Measuring Emotional Responses to Advertising" [18] aimed to improve understanding of the measurement of emotional advertising effects. The study measured "attitude towards the ad" and "attitude towards the brand" using "emotional" attitude scales. When using the "attitude towards the ad" strategy, advertisers designed the ad so as to not directly discuss product attributes but instead leave the consumer in a positive emotional state after viewing the commercial. Ideally, this would create a favorable impression of the ad that would spread to the product. When using the "attitude towards the brand" strategy, advertisers attempted to shape favorable impressions of the brand itself by suggesting favorable consequences of purchasing the brand.

This study was executed as follows: after viewing a set of commercials while watching a television program, subjects were given four minutes and asked to generate verbal response protocols by providing any and all responses to the ads they had just seen. Also, a series of forty-five adjectives was taken from Leavitt's [33] factor analytic study that identified eight factors − amusing, authoritative, dislike, energetic, familiar, novel, personal relevance and sensual − to rate television commercials. Each factor consisted of six adjectives having the highest loadings on that factor. Respondents were asked to judge on a five point-scale whether each of the adjectives described the commercial "extremely well" (5) to "not very well at all" (1). We use this framework in our study to evaluate and label ads.

2.2 Social media used to help measure business impact

As mentioned above, there is a great deal of interest in this intersection of social media and advertising. While our work will build on the work recently done in extracting value from social media, it also relates to more traditional advertising work. We will build on the TV advertising [15-21] literature, particularly work that focuses on Super Bowl TV advertising.

Social media is also being used to enhance the impact of customer service organizations and help customer service representatives tune into the likes and dislikes, praises and complaints customers have about particular products through their social media commentary. Companies like Attensity boast of "going beyond basic sentiment to get detailed reports on customer feedback about new products, campaigns, brands, service and support, and other business drivers" [34].

Social media analysis has the potential to help brands identify opportunities, threats, and assess the need or ways in which they can protect their brand reputation. In the realm of public relations, blogs and social media have driven public relations into the direction of facilitating more two-way communication by opening up direct channels of communications between organizations and their publics and providing opportunities for public relations practitioners to build relationships with certain target publics [35].

2.3 Current commercial measuring of social media response in advertising

Most of the companies measuring social media response to television programs and or commercials generally use methods to analyze sentiment and trends. Sentiment analysis is a set of methods implemented in computer software, that detect, measure, report, and exploit attitudes, opinions, and emotions in online, social, and enterprise information sources [36]. Basic sentiment analysis looks at positive, negative, or neutral comments being generated; however, more complex sentiment analysis can be applied measuring vulgar or polite writing style, serious or amused, calm or excited. Trending analysis looks at the volume of comments taken during a specific timeframe, during a show or commercial, and maps it to actual events during the show or commercial. This gives an indication of which scenes/ characters/events influence social media commentary.

Our study differs from current commercial efforts in that we seek to measure advertisement using real-time social media content analysis to improve our understanding of the factors contributing to specific types of social media response along a large number of TV advertising dimensions. We link the traditional measure of advertising through commercials to the content of what is being said through social media.

What we care about is measuring the effectiveness of ads by peeking at what people were saying at the time of the advertisement. We seek to show how this creates a better understanding of the effectiveness of the advertisement using gender-based response, and popularity metrics beyond just the frequency of response like the number of tweets about the actual product. We also conduct deeper analysis into whether people were talking about the product throughout the whole game (not just during and right after the commercial) and, most importantly, determining if there was an increase in the number of followers of a product after the ad. These latter metrics reflect increased sustained interest in the product. By not just looking at popularity, but we can give a better analysis by looking deeper into the content of what is being said.

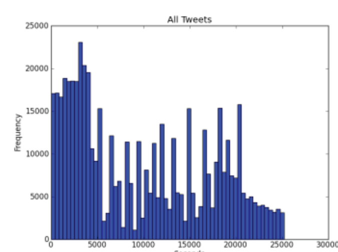

Figure 1. Distribution of tweets in our database over the duration of the game. Time zero corresponds to 5pm February 5th, 2012. The spikes correspond to Super Bowl events such as touchdowns. There is more online activity in the beginning of the Super Bowl than in the end.

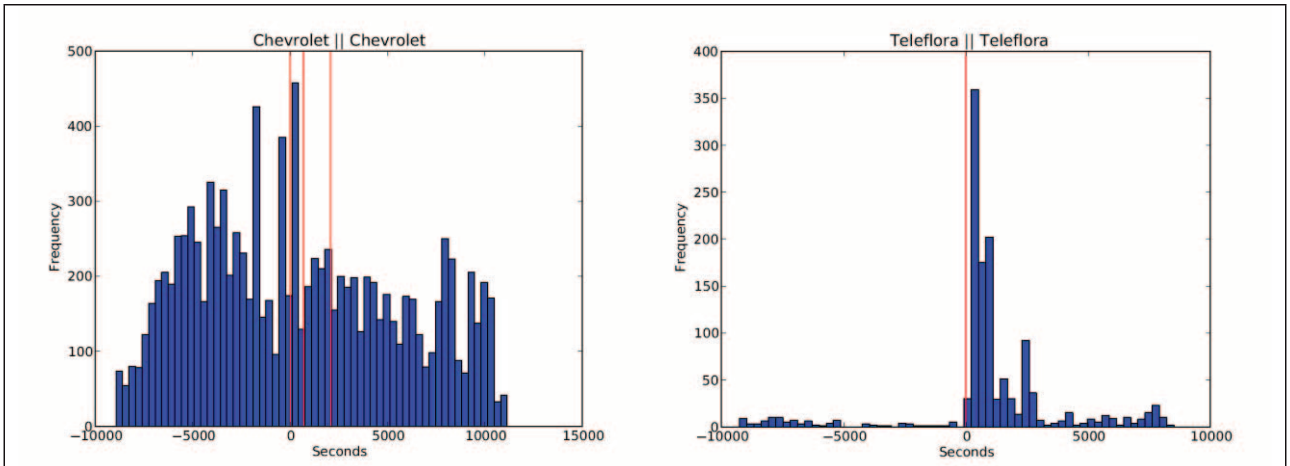

Figure 2: Frequency of social media response over time to Super Bowl Advertisements for Chevy (left) and Teleflora (right). Vertical line indicates time the first ad for the brand was shown. The horizontal axis is the number of seconds before and after the commercial airing. The vertical axis is the Frequency of the number of mentions in our database.

3. Data Description

We use three types of data in this study: 1) evaluations of commercials by raters (discussed in the Method section below); 2) time-stamped recordings of the Super Bowl advertisements (Our research team labeled events that happened during the Super Bowl to synchronize those events with what social media users were discussing online.); and 3) tweets.

The Super Bowl aired from 6:30 pm EST until approximately 10 pm EST on February 5th. The Twitter data was collected from 5pm EST until 12am EST on February 5th 2012. The excess time on both ends allowed for the ability to estimate baselines for buzz baseline effect of how people were tweeting about the advertisers before the actual Super Bowl began. The collected Twitter data consisted of a set of records where each record consisted of fields associated with the Twitter message–the most important of which were user name (a field where a user could give a name), a 140 character text string (tweet), geographic latitude and longitude (optional), and a timestamp of when the tweet occurred. The usernames were matched to a gender list and then were scrubbed (or thrown away) before storing the tweets into our database. There were 534,482 records, which came from 406,031 distinct users. Roughly 1.4% of these records contained geographic data. Figure 1 shows the distribution of these records over time, where 0 is the start of the collection of Twitter data at 5PM EST, lasting 25,200 seconds until 12AM EST. The peaks in the data tend to correspond to major moments in the Super Bowl game (i.e. points being scored, lead changes, presentation of the MVP trophy).

The data was collected using the Twitter Streaming API, which allows a user to submit a list of keywords and retrieve a subset of all publicly available tweets that contain those keywords posted in real-time, with a maximum of four thousand per second. This means that for some popular keywords, a user may not retrieve all tweets. The list of keywords consisted of the brands and celebrities that were going to be in the commercials. The list of commercials was publicly available beforehand and we used this to set up our queries in advance.

In preliminary data analysis, we found that social media response in terms of volume is vastly different across advertisements. Response is almost instantaneous and often short-lived on Twitter. On first inspection, we find the shape of the distribution of buzz response varies by product and product category. In Figure 2, we show the social media response over time to two advertisements, Chevrolet and Teleflora. The figures indicate how varied the responses can be. The left-most plot shows a campaign that resulted in buzz throughout the Super Bowl. The right-most plot shows a campaign that resulted in a lot of buzz, but only in the minutes during and shortly after the advertisement. The horizontal axis reflects the time in seconds. Time 0 corresponds to the time the first ad for the brand was shown. The vertical axis is the frequency of the number of tweets in our database. The red vertical lines indicate the times the ads were shown.

In addition to the preliminary analysis of number of tweets, we analyzed the tweets by gender response and the number of new followers the brand garnered. Table 2 shows the proportion of female, male and unknown gender tweeters by advertisement. The gender of a user was inferred via a male/female dictionary match. Tweeters whose gender could not be inferred based on exact name match to a list of female and male names were labeled as unknown.

In addition to the amount of buzz, we were able to monitor the increase in followers of the brands. Table 3 shows the Twitter brands associated with the Super Bowl and the amount of increase in following they received during the Super Bowl. They are ranked by increased proportion of followers on the left and absolute number increase on the right. Note that while Chevrolet was not the most popular ad in terms of absolute number of tweets, it was the brand that received the most new followers during the Super Bowl.

The aforementioned examples illustrate the richness and potential of the data. These data discussed above were used to construct features to be used in a linear regression model. The list of features constructed and modeling approach is discussed below.

4. Method

In this paper, we aim to identify specific ad features and social media campaign strategies that lead to different types of social media outcomes advertisers may care about. Those outcomes include volume (or the amount of social media buzz), valence (positive, negative and neutral sentiment), and reach (a measure of how sub-communities are engaged). We use linear regression to relate a number of advertising and social media related attributes to advertising outcomes in real time. The attributes used in our study are discussed below.

Table 2: Proportion of female/male responders by advertiser ranked in descending order by female response

Ad	FemaleProp	MaleProp	UnknownProp
NBC-The Voice	0.57	0.12	0.30
NBC-Smash	0.47	0.21	0.33
Careerbuilder	0.45	0.20	0.35
Audi	0.45	0.28	0.27
Teleflora	0.45	0.20	0.35
Best Buy	0.45	0.27	0.29
Volkswagon	0.43	0.31	0.26
Sketchers	0.43	0.28	0.30
H&M	0.42	0.18	0.41
Kia	0.41	0.29	0.29
Pepsi	0.41	0.21	0.38
Movie-John Carter	0.40	0.30	0.30
Oikos Greek Yogurt	0.39	0.23	0.39
M&M	0.38	0.24	0.37
NBC	0.38	0.31	0.30
Doritos	0.38	0.20	0.42
Honda	0.38	0.34	0.28
Etrade	0.38	0.28	0.34
Lexus 320GS	0.38	0.38	0.25
Verizon	0.37	0.17	0.46
Pepsi	0.37	0.30	0.33
Go Daddy	0.37	0.37	0.26
History Channel	0.36	0.37	0.27
cars_com	0.34	0.32	0.35
GE	0.34	0.30	0.36
Grand Total	0.33	0.31	0.35
Superbowl	0.33	0.31	0.36
Toyota	0.33	0.27	0.40
Chevrolet	0.33	0.34	0.33
Hulu	0.33	0.38	0.29
Budweiser	0.33	0.38	0.30
Movie-Avengers	0.32	0.34	0.33
Coca-Cola	0.31	0.32	0.37
Century 21	0.30	0.34	0.35
Fiat 500	0.30	0.37	0.34
Movie-Battleship	0.30	0.33	0.37
Movie-Act of Valor	0.30	0.42	0.28
Chyrsler	0.29	0.43	0.28
MetLife	0.29	0.42	0.29
Cadillac	0.28	0.38	0.34
Movie-GI Joe	0.27	0.36	0.37
Hyndai	0.27	0.34	0.39
TaxAct	0.27	0.44	0.29
Movie-The Lorax	0.27	0.27	0.47
Bridgestone	0.25	0.34	0.42
Acura	0.24	0.32	0.44
Ford	0.16	0.29	0.56

4.1 Advertiser attributes

Forty-six brands advertised during Super Bowl 2012. A large number of these advertisements were posted on YouTube. So we had many raters evaluate the creative aspects of the YouTube commercials. We used the factors in Leavitt's [33] factor analytic study that identified eight factors−amusing, authoritative, dislike, energetic, familiar, novel, personal relevance, and sensual−to rate television commercials. We had 60 responders rate each commercial. The ratings on these factors are used as independent variables in our model. In addition to the aforementioned Advertising characteristics, we constructed two other features about the commercials: whether or not the commercial included a celebrity or an animal. We also constructed some controls: the number of ads for the brand during the Super Bowl, the length of the first commercial for the brand in seconds, the number of followers on Twitter the brand had before the Super Bowl, and the number of tweet posts for the brand from 5pm-5:05pm. This level of buzz at 5pm was used as a baseline measure of "buzz."

4.2 Social media attributes

We are primarily interested in how the social media strategies of the firm lead to more or less buzz, more or less followers, more or less positive sentiment. The three types of social media strategies we labeled were: 1) whether the brand provided a social media game to engage viewers during the Super Bowl; 2) whether the brand included a link to a social media site (example, Twitter or Facebook page) in the advertisement; and 3) the number of Tweets the brand made during the Super Bowl.

4.3 Dependent Variables

For dependent variables or outcomes, we studied the amount of buzz within the five minutes after the ad, the half-life of the ad (which we define as the relative change in buzz five to ten minutes after the ad aired compared to the buzz about the brand an hour-and-a-half before the Super Bowl began), the increase in Twitter following, the sentiment (we built our own sentiment scorer for tweets), the proportion of female followers, and how many tweets were about the product as opposed to the creative aspects of the ads. Table 4 lists these features.

Because our data set is small, representing only a few brands and we have a large number of predictors, we can only claim the results as preliminary and make loose claims, there is still work to do to verify our results. However, in the results section below, we discuss the preliminary findings of the significant social media variables linked to outcomes that firms care about. We consider six dependent variables. The results of the linear regression models are discussed below.

5. Results

In this section we present a discussion of the attribute subsets that were found to be the most correlated with the different dependent variables of interest. In Table 5, we present the results of models built on the social media features controlling for the size of the Twitter network before the Super Bowl. We present two R-sq values for each dependent variable, the first is based on a model built on all of the attributes discussed above and the second is the R-sq value of the model built on only the social media features and the network control. We exclude Movies from the results in Table5 because they appear to behave differently than the other types of brands with respect to social media response.

Our preliminary results reveal that indeed, different factors associated with TV advertising seem to matter more for different dependent variable outcomes. The number of tweets is highly correlated with the emotional characteristics of the ads and whether there was social media in the ads, while the proportion of tweets about the product is highly correlated with product type. The sentiment is highly correlated with the type of brand and the creative aspects of the ads. Still, the NetImp10 and NumTweets models indicated that having a social media strategy is a significant link to social media success.

Table 3: Twitter follower increase for handles associated with the Super Bowl. The left column shows the proportionate increase and the right shows the absolute number of new followers during the Super Bowl. Chevrolet was the winning brand by increased absolute number of followers. Chevy was posting on Twitter during the Superbowl and held a social media competition.

Handle	Percent	Handle	Number
M&Ms	14.70%	Chevrolet	1852
Act of Valor	12.29%	Avengers	1823
Battleship	7.60%	NBC The Voice	1552
GI Joe	6.90%	Audi	1492
Avengers	6.85%	Volkswagon	1097
Chevrolet	2.56%	Coca-Cola	1056
John Carter	2.42%	H&M	949
Volkswagon	1.92%	The Hunger Games	852
Bud Light	1.46%	GoDaddy.com	364
Doritos	1.37%	M&Ms	307
The Real Dictator	1.33%	Act of Valor	285
Teleflora	1.20%	Pepsi	257
Acura	0.85%	History Channel	150
NBC The Voice	0.80%	Doritos	145
Audi	0.75%	Battleship	141
Chrysler	0.73%	Toyota	140
Pepsi	0.61%	Fiat	135
The Hunger Games	0.59%	Chrysler	134
Hyundai	0.38%	General Electric	110
Bridgestone	0.37%	GI Joe	109
RapSheet	0.37%	Acura	91
Fiat	0.35%	Hyundai	83
Met Life	0.34%	Best Buy	76
General Electric	0.33%	John Carter	75
GoDaddy.com	0.30%	Bridgestone	71
Coca-Cola	0.22%	Honda	70
Honda	0.20%	Lexus	53
Cadillac	0.19%	Teleflora	49
Toyota	0.19%	Samsung	47
E*TRADE	0.18%	Cadillac	45
Kia	0.18%	Hulu	39
Taxact.com	0.16%	Bud Light	33
H&M	0.12%	CareerBuilder.com	30
Cars.com	0.12%	Kia	20
History Channel	0.08%	Century 21	19
Hulu	0.06%	Taxact	8
CareerBuilder.com	0.05%	Cars.com	8
Pepsi	0.04%	The Negotiator	6
Best Buy	0.03%	Dannon	4
Samsung	0.03%	The Real Dictator	3
Dannon	0.02%	Met Life	3
Lexus	0.02%	E*Trade	3

Our findings that social media strategy leads to success with increased followers agree with the fact that Chevrolet, a brand that had a solid social media strategy including a competition during the Super Bowl, led with increased following while it and was a popular ad in general with respect to volume. Finally, whether an ad was tweeted by female users seemed to be most correlated with the brand type.

Although our data are small, our findings are intuitive and illustrate that the number of tweets seems to be more correlated with the creative aspects of ads than the proportion of tweets about the product features, which appears to be more correlated with the type of product. Further, the number of increased followers can be significantly linked to social media strategy.

It is not surprising that product type matters for online word of mouth. Our findings are consistent with the word-of-moth literature in Marketing. In our data, product type also seems to matter for increasing number of followers–users are more likely to follow movies than cars, and more likely to follow cars than general technology products, for example. See Figure 3 for outcome variables by brand type. That said, still, the volume of buzz seems to be driven primarily by the creative aspects of ads and not just the product type, indicating that advertisers might need to create advertisements that get people talking about the product. The main take-away of this work is that whether a brand had a social media strategy is very important when it comes to increased followers. The punch line therefore, is that social media in the ad seems to be linked to buzz and other advertising outcomes, at least for Super Bowl 2012.

Table 4: Variable description

Type	Name
Control	**Type:** Type of product/brand, **Qtr**: The Quarter the first ad for the brand was shown, **NumAds:** Number of ads during the Super Bowl for the brand, **Length**: Length of the ad, **Net4pm:** The size of the follower network on Twitter at 4pm, **Baseline5pm:** The number of tweets about the brand in our database between 5pm and 5:05pm
Hypotheses (attributes of interest)	Rating on Leavitt's eight attribute dimensions of the creative **Amusing, Authoritative, Dislike, Energetic, Familiar, Novel, PersonalRelevance, Sensual. Celeb:** Was there a celeb in the ad or not? **Animal:** Was there an animal in the ad or not? **SocialMediaGame:** Did the brand provide a social media game to engage viewers during the Super Bowl? **SocialMediaIntheAd:** Did the brand include a link to social media site (example, Twitter or Facebook page) in the advertisement? **CommOnSocialMedia:** The number of Tweets the brand made during the Super Bowl.
Dependent	**Netimp10:** Proportion of network growth at approximately 10pm, just after the Super Bowl, **Half-life:** Proportion of Twitter follower growth in the 5-10 minute window after the ad aired to the 5-10 minute baseline window from 5pm-5:05pm EST Feb 5, 2012, **Product:** Proportion of tweets about the features of the product, **NumTweets:** number of tweets about the brand in our database within five minutes of the ad being aired, **Sentiment:** The average sentiment of the posts about the ad, **Female response:** The proportion of female responders to the ad

6. Limitations

Given that the dataset we analyze is observational, there are many events over which are difficult to control for. The data are extremely small and therefore we had to limit the features in the models that we presented . In the future, we will evaluate multiple commercials in multiple TV settings. We have already collected social media response to advertisements during the very popular US TV show finales, The Voice, Dancing with the Stars and American Idol finales.. Despite having response to more advertising events, the data will still be small and may limit our ability to make any causal claims. However, first steps to link social media strategy to buzz outcomes are very important. Nevertheless, in this paper, we recognized the Super Bowl is an extreme case where TV viewers are focused on the ads and are therefore more likely to discuss them. In the future, we will also look at response on other social media outlets (for example, Facebook). Finally, data collection was time consuming. We hope to automate the process of timing commercials. We are in the process of testing tools for this.

7. Discussion and Conclusions

The main contribution of this paper is a look at the content of social media buzz in response to TV advertising in real-time – as opposed to, say, only evaluating response by popularity or sentiment. We believe it is the first time that researchers have

	Netimp10	Half-life	Product	NumTweets	Sentiment	Female
Intercept	-958.765*	-0.054	-0.218	35.040	16.924	0.216
SocialMediaGame	277.134+	0.0393+	-0.013	504.402*	12.137	0.004
SocialMediaIntheAd	241.334	-0.182	0.105	-14.570	-0.445*	0.041
CommOnSocialMedia	5.118*	-0.007*	-0.001	-5.056+	-0.159	0.030
Net4pm	223.136*	0.1781	0.0677	87.044*	-0.197	0.029+
N	30	30	30	30	30	30
R-sq with just social media attributes and Net4pm	0.429	0.4869	0.1787	0.192	0.184	0.208
R-sq overall	0.669	0.565	0.420	0.677	0.419	0.649

+p<.1, *p < 0.05, ** p <.01, *** p <0.001

Table 5: Social Media Factors correlated with the outcomes

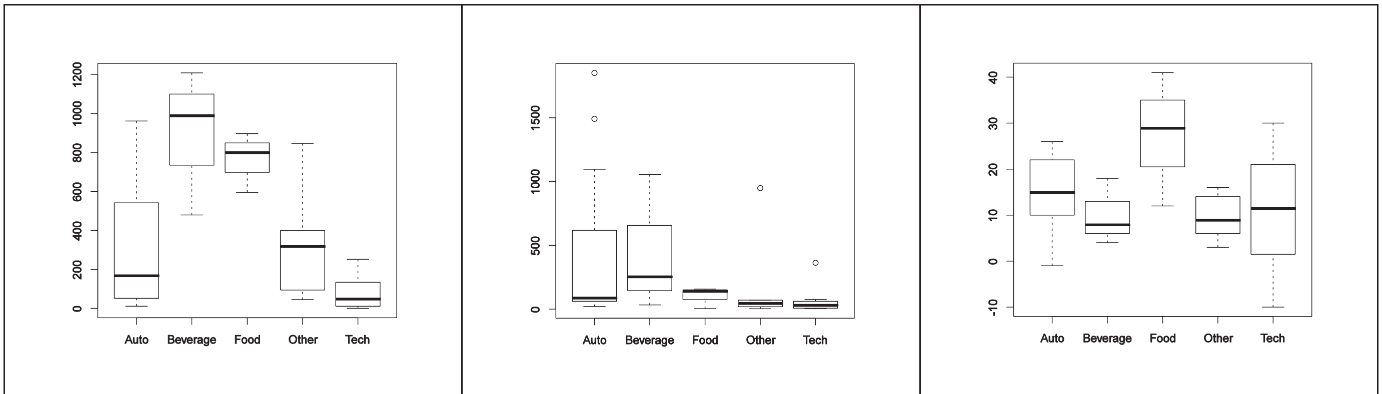

Figure 3: Outcome variables by product type. Number of Tweets (left), Number of New followers (center), Sentiment (right).

tried to link content in ads to content of discussions on the web in real-time. Currently there is no set industry standard for measuring social media success. Various metrics such as increased fans or followers, increased brand recognition, and increased sales could be used to help define and measure the business impact of social media. In this paper we study how specific features of ads and social media strategies lead to specific types of responses, in particular response about the product itself, increased following, sentiment and overall volume; a first step at looking at how the content of online social media buzz varies based on the stimulus–in this case, different types of ads covering different types of brands.

Certain factors "predict" buzz better, while others predict followers and mentions about the product and who is likely to respond positively better. By understanding the content of buzz, advertisers can better respond to trends and develop product following, ultimately leading to increased sales and a better return on investment for advertising dollars. Furthermore, there is the potential for advertisers not only to listen passively to consumers, but to develop strategies to engage directly with consumers as was the case with the brands that ran competitions. In the future, social media will be used to make better advertisement sales decisions in terms of placement and content. Social media analysis can indicate which characters, storylines, and so on have the most impact on the target audience and can also provide real time information on who is talking about a particular ad (stats about social media commentators such as gender, location, and even lifestyle).

To illustrate the use of social media to identify response within different groups, we measured the overall sentiment of female and male tweet posters over products that were mentioned during the super bowl. Table 6 lists the average sentiment of tweets posted by each of these groups. From this table, it is clear that female sentiment regarding a product is typically greater than male sentiment (possibly due to more polite language).

In order to explore these differences more deeply, we examined a subset of male/female Samsung-related tweets to better understand this difference in average sentiment. A sample of low sentiment male tweets are listed below (expletives removed).

- rt @name: dear #samsung, no one wants a f***ing stylus.. not even android fanboys − sincerely, the internet.
- i liked the knock on apple re: samsung but are they just relying on the tablet having a pen/stylus to sell? lame. so palm pilot circa 2000
- good: giants win bad: patriots intentionally letting their opponents score ugly: that horrid samsung commercial for a phone with a stylus

A selection of positive tweets posted by female users are listed below.

- @name @todayshow palm pilot won. that #samsung ad felt retro. pen?
- rt @name: high praise indeed-- thanks for the shout-out rt @name the darkness ad is my fave of the night. #samsung ...

There were also some cases where female users had a lower average sentiment than male users for a brand. Some of these brands (e.g., Teleflora) had a greater number of female tweets than male tweets, indicating that raw buzz may not be enough to capture public engagement. Two striking examples are related to the Kia and Teleflora advertisement shown. In these examples, the average sentiment for female posts was 0.027 and 0.024 lower than the average sentiment for male posts, respectively. Some examples of strongly negative female tweets, and strongly positive male tweets are below in Table 7 below.

Table 7: Examples of positive and negative tweets by gender

Female tweets	Male tweets
rt @name: kia sexually exploits women too in their ads. #superbowl	best #superbowl commercial: kia with the crue. because the crue, adriana lima, girls in bikinis, race cars, giant sandwich, happy ending.
the worst thing about this kia optima commercial is that i thought it was a great looking car.....until this commercial #cheapified	love the kia optima commercial. very funny... a new classic. motley crue & kia - a match made in heaven?
rt @name: i m not a f***ing commodity, @teleflora. no one should be coerced into sex. #notbuyingit	i m calling it. and the winner of this #superbowl is... @adrianalima for teleflora. not only was it the best, it was sexy & super cheeky.
@teleflora that super bowl ad was disgusting and demeaning to women. i would expect that from godaddy, not you.	@adrianalima @teleflora hey adriana you looked very attractive in your commercial.

From these tweets, it is clear that social media users not only discuss the creative aspects of advertisements, but also broadcast very passionate opinions on the products themselves.

Social media analysis can help brands identify opportunities, threats, and assess the need or ways in which they can protect their reputations. Our work is a first step towards understanding the link between what is said on TV and what is said in response on social media. We also show in our current and ongoing work that using real time social media response is a new way to identify social network effects [37] .
We have collected an incredibly rich dataset, and have only scratched the surface of what it offers. We also intend to make the data publicly available for the research community eventually. In terms of future modeling, we will consider using survival models to measure the amount of social media buzz about ads. The benefit will be that we can increase our sample size due to the many data points available per ad over time. We believe our methods, once fully developed, will have broader implications outside the context of advertising.

Table 6: Male/female sentiment regarding various advertised products/brands. This only includes the top and bottom five brands with greatest difference between male and female sentiment.

Brand	Male		Female	
	avg sentiment	count	avg sentiment	count
Verizon	0.66	2167	0.76	1574
Samsung	0.66	2147	0.76	1567
Lexus	0.66	149	0.72	134
Budweiser	0.61	4041	0.66	2395
Honda	0.74	768	0.80	839
Fiat	0.68	1462	0.66	662
Cars.com	0.68	339	0.66	189
Teleflora	0.76	232	0.74	368
Sketchers	0.79	140	0.76	97
Kia	0.80	1042	0.77	703

One trending area in which Social TV can have an impact is in election campaigns. Using real-time analytics of social media generated during the GOP primary debates, Yahoo and the social analytics firm Attensity were able to provide analytics of the candidates' performances during the debates [38]. With the success in this endeavor, it can be expected that Social TV will play a significant role in the discourse during the 2012 Presidential debates. Using methods for content analysis, political strategists and campaigns can better streamline their messages to fit their audiences and, in turn, instigate specific responses.

8. Acknowledgements

The data collection for this project was labor intensive and the final dataset including timed commercials, tweets and scoring of commercials was the product of the work of many research assistants. We would like to thank Anthony Crawford, Michelle Gittens, John Paul Lacovara and Haftom Khasai for their help. In addition, we would like to thank Rob Malcolm and Aretta Baldon for valuable comments about norms in advertising measurement in the real world. Related work can be found on our lab page (www.thesocialtvlab.com).

REFERENCES

1. Dumenco, S. *7 Things You Need to Know About 'Social TV' Right Now The State of the Art -- and Science -- of Monitoring Social-Media Chatter About TV*. 2011 [cited; Available from: http://adage.com/article/the-media-guy/social-tv/229845/.
2. Talbot, D. *A Social Media Decoder*. 2011 [cited; Available from: http://static.bluefinlabs.com/website/bluefin_mit-tech-review.pdf.
3. Proulx, M. and S. Shepatin, *Social TV: How Marketers Can Reach and Engage Audiences by Connecting Television to the Web, Social Media, and Mobile*. 2012: Wiley.
4. Beyers, D. *111.3 million people watched Super Bowl* 2012 [cited; Available from:

http://www.politico.com/blogs/media/2012/02/million-people-watched-super-bowl-113632.html.
5. Ghose, A., P.G. Ipeirotis, and B. Li, *Designing Ranking Systems for Hotels on Travel Search Engines by Mining User-Generated and Crowd-Sourced Content.* Working Paper, 2011.
6. Berger, J. and K.L. Milkman, *What Makes online Content Viral?* Journal of Marketing Research (Forthcoming), 2012: p. 1-17.
7. Berger, J. and E.M. Schwartz, *What Drives Immediate and Ongoing Word of Mouth?* Journal of Marketing Research, 2011. **48**(5): p. 869-880.
8. Berger, J.A. and R. Iyengar, *How Interest Shapes Word-of-Mouth Over Different Channels.* Working Paper, 2011.
9. Hill, S., F. Provost, and C. Volinsky, *Network-Based Marketing: Identifying Likely Adopters via Consumer Networks.* Statistical Science, 2006. **21**(2): p. 256-276.
10. Chen, H., P. De, and Y.J. Hu, *IT-enabled broadcasting in social media: An empirical study of artists' activities and music sales.* Working Paper, 2011.
11. Dhar, V., III and E. Chang, *Does Chatter Matter? the Impact of User-Generated Content on Music Sales.* Journal of Interactive Marketing 2007. **23**(4): p. 300-307.
12. Benton, A., et al., *A system for de-identifying medical message board text.* BMC Bioinformatics, 2011. **12**(Suppl 3): p. S2.
13. Benton, A., et al., *Identifying potential adverse effects using the web: A new approach to medical hypothesis generation.* Journal of Biomedical Informatics, 2011. **44**(6): p. 989-996.
14. Peres, R., Shachar, Ron and Lovett, Mitchell J, *On Brands and Word-of-Mouth*, in *Available at SSRN: http://ssrn.com/abstract=1968602 or http://dx.doi.org/10.2139/ssrn.1968602*. 2012.
15. Fam, K.-S., *Attributes of Likeable Television Commercials in Asia.* Journal of Advertising Research, 2008. **48**(3): p. 418-432.
16. Farrall, N. and J. Whitelock, *A comparative analysis of advertising characteristics, strategy, style and form in global and national brand advertising.* Journal of Marketing Communications, 2001. **7**(3): p. 125-136.
17. Fehle, F., S. Tsyplakov, and V. Zdorovtsov, *Can Companies Influence Investor Behaviour through Advertising? Super Bowl Commercials and Stock Returns.* European Financial Management, 2005. **11**(5): p. 625-647.
18. Hill, R.P. and M.B. Mazis, *Measuring Emotional Responses to Advertising* Advances in Consumer Research, 1986. **13**: p. 164-169.
19. Moorman, M., P.C. Neijens, and E.G. Smit, *The effects of program-involvement on commercial exposure and recall in a naturalistic setting.* Journal of Advertising, 2007. **36**(1): p. 125 - 141.
20. Smit, E.G., P.C. Neijens, and L.v. Meurs, *Effects of Advertising Likeability: A 10-Year Perspective.* Journal of Advertising Research, 2006. **46**(1): p. 73-83.
21. Till, B.D. and D.W. Baack, *Recall and Persuasion: Does Creative Advertising Matter?* Journal of Advertising, 2005. **34**(3): p. 47-57.
22. Tomkovick, C., R. Yelkur, and L. Christians, *The USA's biggest marketing event keeps getting bigger: an in-depth look at Super Bowl advertising in the 1990s.* Journal of Marketing Communications, 2001. **7**(2): p. 89-108.
23. Aral, S., Ipeirotis, P, Taylor, S., *Cramer's Rule: How Information Content Moves Markets*, in *Workshop on Information Systems Economics*. 2009: Phoenix.
24. Horovitz, B. *Super Bowl ads grab social-media buzz* 2012 [cited; Available from: http://www.usatoday.com/money/advertising/story/2012-02-06/super-bowl-ad-panel-favorites/52981424/1.
25. Elliot, S. *Super Bowl Commercials Rated by Social Media* 2012 [cited; Available from: http://www.nytimes.com/2012/02/08/business/media/super-bowl-commercials-rated-by-social-media.html.
26. Grinberg, M. *Superbowl 2012 Breaks Social Records, Sets New Marketing Standard*. 2012 [cited; Available from: http://www.huffingtonpost.com/maya-grinberg/superbowl-2012-breaks-soc_b_1260869.html.
27. Kellogg. *Super Bowl Advertising Review*. 2012 [cited; Available from: http://www.kellogg.northwestern.edu/news/superbowl/index.htm.
28. Wharton. *Wharton Future of Advertising Super Bowl Ad Tweet Meet 2012*. 2012 [cited; Available from: http://myfoa.net/superbowlads/info/.
29. Terry-Mcelrath, Y., et al., *The Effect of Antismoking Advertisement Executional Characteristics on Youth Comprehension, Appraisal, Recall, and Engagement.* Journal of Health Communication, 2005. **10**(2): p. 127-143.
30. Kover, A.J., S.M. Goldberg, and W.L. James, *Creativity Vs. Effectiveness? An Integrating Classification for Advertising.* Journal of Advertising Research, 1995. **35**(9): p. 29–39.
31. Ang, S.H. and S.Y.M. Low, *Exploring the Dimensions of Ad Creativity.* Psychology and Marketing, 2000. **17**(10): p. 835-854.
32. Stone, G., D. Besser, and L.E. Lewis, *Recall, Liking, and Creativity in TV Commercials: A New Approach.* Journal of Advertising Research, 2000. **40** (3): p. 7–18.
33. Leavitt, C., *A Multidimensional Set of Rating Scales for Television Commercials.* Journal of Applied Psychology, 1970. **54**: p. 427-429.
34. Attensity. *Social Analytics: Go beyond social media monitoring with Real-Time Social Analytics*. 2012 [cited; Available from: http://www.attensity.com/attensity-solutions/business-solutions/social-analytics/.
35. Wright, D.K. and M.D. Hinson, *How blogs and social media are changing public relations and the way it is practiced.* Public Relations Journal, 2008. **2**(2): p. 1-21.
36. Helweh, A. *The Future of Sentiment Analysis* 2011 [cited; Available from: http://www.socialmediaexplorer.com/social-media-monitoring/sentiment-analysis/.
37. Hill, S., Benton, A. *Social TV: Analyzing the Impact of Social Media Strategies on Viewer Engagement.* in *Workshop on Social Media Analytics (SOMA), KDD.* 2012
38. MacDonald, R. *Attensity and Yahoo! Bring Real-Time Social Analytics to Republican Presidential Debate*. 2012 [cited; Available from: http://www.attensity.com/2012/01/06/attensity-and-yahoo-bring-real-time-social-analytics-to-republican-presidential-debate/.