(When) Can Social Media Buzz Data Replace Traditional Surveys for Sales Forecasting?

Yuying Shi
Texas A&M University-Commerce

Ekaterina V. Karniouchina
Mills College

Can Uslay
Rutgers University

Abstract

Producing reliable sales forecasts for new products is notoriously difficult. Traditional surveys have been popular for decades but they are relatively cumbersome and expensive to implement. Use of readily available data from social networks is becoming increasingly popular. Can the customer buzz measures obtained from social networks be a convenient and inexpensive substitute of traditional tools? Using movie industry as a backdrop, the authors compare forecasts derived from a large-scale purchase intention survey with those obtained via social networks. They find that while the survey approach is more reliable in predicting performance overall, customer buzz-based forecasts outperform surveys under conditions of high uncertainly, e.g., for niche and low-budget movies.

Introduction

Innovation is widely viewed as critical to firms' stability and long-term survival. Firms typically spend more of their resources when commercializing innovation compared to research and development. Many new products fail as managers hesitate to pull the plug on projects before the commercialization stage due to "escalation of commitment." [1] Nevertheless, firms spend significant resources trying to assess the potential of new

products before they enter markets since eliminating unsuccessful projects and reallocating marketing resources toward more promising products can avoid waste, and help maximize performance. However, predicting demand for new products remains one of the biggest challenges that businesses face.[2]

Purchase intention survey

"If one wants to know whether or not an individual will perform a given behavior, the simplest and probably the most efficient thing one can do is to ask the individual…"[3] This famous quote justifies the wide popularity of the survey approach. As such, consumers' stated purchase intention has been commonly used to forecast new and existing product sales for decades.

A typical purchase intention question is: How likely are you to purchase this product? The answer is presented on a 10-point scale with 10 indicating "extremely likely" and 1 indicating "extremely unlikely." Respondents are traditionally asked to select a category that corresponds to their purchase intention. Researchers summarize survey results using a metric such as *Top Box*[4]and/or make other adjustments based on empirical history of working with such data. However, surveys are relatively cumbersome and expensive to implement. The process includes developing a customized survey instrument for a given product/service, identifying the right market segment, recruiting a group of respondents that fits in the targeted market segment, and getting a sufficient number of responses. Each step requires expertise, effort, and financial resources.

Customer buzz

Using customer buzz derived via 'listening' to online consumer activity to predict new product success is becoming increasingly popular. Customer buzz can take on various forms: product reviews, online blog/forum postings, and social network and web search activities. The prediction validity of consumer buzz has been examined by scholars and practitioners.[5,6] For example, Adobe claimed that it achieved a 100% accuracy level when predicting whether a movie's domestic receipts would exceed its production budget using social media data from platforms such as Twitter and Facebook. Inspired by this success, Adobe deemed the use of social media in predictive intelligence highly worthy by calling social media "a giant water cooler that is underutilized."[7] Similarly, a technology journalist reported that the mentions of the new iPhone 8 on Weibo, China's most popular social media platform, were sparser than those for earlier iPhone models. Since China represents the largest Asian market for iPhone, he was able to correctly predict a soft demand of iPhone 8 in Asia market.[8]

Companies are starting to replace traditional insight generation strategies with Big Data Analytics (BDA). If we look at the broader market for BDA, its size is expected to reach 274 billion by 2022.[9] The areas of social media listening and analytics that is of particular importance to this study are rapidly expanding. For example, the global social media analytics market is forecasted to grow at a rapid pace from $3 billion USD in 2019 to $9.4 billion USD in 2024.[10] Companies adjust their marketing strategies and even launch new products in response to the insights obtained from social media data combined with other sources. For instance, Ben and Jerry's noticed an uptick in social media buzz during bad weather, which is counterintuitive since people traditionally associate spikes in ice-cream sales with hot weather. It turns out that people like to stay indoors and watch Netflix while eating ice-cream when the weather is not amenable. Based on these findings Ben and Jerry's did not only begin to incorporate bad weather sales spikes into their advertising strategies (e.g. boosting ad spending during rainy weather and winter storms) but went a step further and launched a new Netflix & Chill'd flavor.[11]

In the era of Big Data, such advantages of utilizing consumer buzz are particularly prominent. There are numerous sources of inexpensive (or free) readily available online buzz data. The data emerges organically making it superior to biases prone to stated likelihood measures obtained via surveys.

Comparing the two approaches

Given the simplicity of the new approach, should the companies abandon the traditional survey research method and concentrate on aggregating and analyzing online data? The answer is not obvious. Some may argue that surveys represent a better approach because they elicit purchase intentions directly. In contrast, the product reviews or the social media mentions are, at most, serve as a proxy for a potential behavioral outcome. Thus, any attribution to sales is more intuitive for the former than the latter.[12] Others may argue that buzz predictions are more accurate because surveys are not always representative since they are prone to biases in sample selection.

For new product predictions, this issue is particularly important because given the large degree of uncertainty associated with many new products, an established segment is hard to identify and the acceptance level of the potential target market can vary dramatically. In addition, the new concepts/features valued by the designers may not be well-received by the market. Customer buzz reflects an aggregation of opinions of a much bigger sample size than the survey. The larger sample size may help to mitigate the representation biases of the survey approach. In addition, potential early adopters will self-select and leave their digital traces in spaces where new

products of interest are featured or discussed while it is notoriously difficult to assemble highly customized samples for the purposes of the survey with many potential participants failing to pass the screening stage or providing answers while being relatively unformed.

In sum, both approaches have relative advantages and disadvantages. Assessing the relative accuracy of the two approaches will provide insight to managers regarding which one to use when it comes to new product forecasting. If buzz-based predictions prove comparable to survey-based findings, buzz measures can be used as inexpensive substitutes for traditional marketing research tools, given their cost and implementation advantages. Meanwhile, if survey achieves better accuracy, we could further reinforce the validity and value of this traditional approach. In addition to their relative comparison, there are other questions that remain unaddressed. For example, are the two approaches substitutes or complements? What factors contribute to the generation of the purchase intention and customer buzz? With this study, we aim to answer these questions.

Study context

Gaining access to purchase intention data with the corresponding product sales is a challenging task. Most sales records are firms' propriety data. In addition, given that many products have a long product life cycle, the timing of purchase decisions needs to be incorporated into the screening questionnaires and the resulting forecast models. Furthermore, it can be challenging to get respondents to provide opinions on many products in a single questionnaire.

In light of these concerns, the movie industry serves as an ideal context for this study for several reasons. First, because each movie can be regarded as a new product or line extension, and the number of products eligible for inclusion in our sample is quite extensive. Second, film-making is considered a risky industry with extremely high product failure rates,[13] where managers often settle for gut feeling or a "wild guess" when it comes to forecasting.[14] At the same time, accurate prediction is crucial for movie distributors and exhibitors who rely on individual movie forecasts to determine the distribution intensity prior to the official release date,[15] and increasingly for marketers who rely on this media for placing their brands and products and provide a significant portion of a movie budgets.[16,17] Third, movies have a short life cycle. The majority of sales is made in the first few weeks after launch,[18] and the opening-week revenue is an important performance measure of box-office sales. It is not necessary to wait for years for a product to exhaust the majority of its sales. Buzz is a key success driver of the early

adoption of a product with exponentially decaying lifecycle. Movies, as well as media and fashion products, fall into this category.

We conduct a large-scale purchase intention survey on movie going behavior. Customer buzz data are collected through YouTube. We use aggregate metrics to summarize the data and incorporate these metrics in our prediction model. Our primary finding is that survey is better than consumer buzz in its overall performance, but buzz is better than survey approach in predicting performance of specialized products. Both approaches are predictive of movie revenues and they complement each other.

The remainder of this article is structured as follows. First, we introduce our data and methodology. Second, we present our empirical findings. We conclude with a discussion of managerial and research implications.

Data and methodology

We collect all data within two weeks before a movie was officially released because a new product prediction should be conducted before the official release to remove the impact of early adopters and word of mouth (WOM) effects.[19] This way the prediction is based on limited product information and mimics the reality of new product forecasting.

The survey data are obtained using a well-established web lab in a large public research university in the United States. The data collection was conducted weekly for movies widely released between June 2012 and March 2013. We conducted multiple rounds of data collection coinciding with the release dates. The data collection took one and a half years and yielded data on 124 movies. Approximately 1,500 students participated in this study.

Each respondent watched 10-15 trailers of then upcoming movies. For each trailer, the respondent indicated his/her purchase intention by answering the following question: how likely are you to go to the theatre to watch this movie? This question was measured on an eleven-point scale with zero indicating 'very unlikely' to 10 'very likely'. We randomized the order of the trailers as well as the survey questions to minimize presentation order effects.[20] We also collected other information at the end of the session to check for other biases. For example, we asked respondents whether they had already seen the trailer. Subsequent analysis revealed that the responses of subjects who had previously seen the trailer and those who had not were similar.

The Motion Picture Association of America (MPAA) film ratings have G, PG, PG-13, R, NC-17 and Unrated (UR). To be consistent with the university protocol, we did not include NC-17 (Adults only) and UR rated movies. As the research was conducted on a university campus, our sample was younger

than the general movie going population. Therefore, we weighted the sample based on the age category of respondents to match the statistics obtained from the official MPAA report.[21]

Metrics

For purchase intention data, we compute three metrics: *Top Box, Top 2 box* and *Mean* metric (see Table 1). Because buzz reflects market-level information, buzz measures are usually analyzed at the aggregate level. For example, in the previous Adobe example, Adobe counts the number of times a movie's trailer is viewed and the number of times the title was mentioned in social media. We adopt a similar approach. We located the official set of YouTube trailers for each individual film. We compiled three metrics: the number of views, the number of people who liked the movie, and the number of total comments. For movies that have more than one official trailer, we aggregated the number of views for each trailer.

Table 1. Descriptive statistics of various metrics

Metric	Description	Mean	Std. Deviation
Top Box	The percentage of people choosing the highest category	8.49	8.23
Top 2 Box	The percentage of people choosing the two highest categories	13.5	13.97
Mean	Average rating across all respondents	5.33	1.31
# of Views	The number of people who view a movie trailer in YouTube	2.93M	5.08M
# of Likes	The number of people who like a movie trailer in YouTube	11900	17282
# of Comments	The number of people who comment on a movie trailer in YouTube	5650	10418

Note: M stands for million

There is a positive correlation among the two types of metrics (see Table 2). And both types of metrics are positively correlated with the movie

revenue, which means movies receiving higher purchase intention or higher customer buzz are associated with higher revenues.

Table 2. Correlations between metrics

	Top Box	Top 2 Box	Mean	# of Views	# of Likes	# of Comments
Top Box	1					
Top 2 Box	0.964	1				
Mean	0.754	0.818	1			
# of Views	0.481	0.508	0.345	1		
# of Likes	0.507	0.523	0.346	0.947	1	
# of Comments	0.553	0.540	0.353	0.817	0.888	1

Note: All correlation coefficients are significant with p value <= .001.

One interesting thing is that it is very difficult to score high on *Top Box*. Our movie data sets contain several popular movies (e.g., *The Dark Knight Arise*, *Amazing Spider Man*). Although these movies can easily get millions of views in YouTube, none of them gets a *Top Box* over 50%. The highest *Top Box* and *top 2 box* are 48.5% and 61.6% respectively, for the movie *The Dark Knight Arise*.

In fact, any movie that receives a *Top Box* over 20% is already in the 90th percentile for *Top Box* score among all movies in our sample. Figure 1 presents all 12 movies with a *Top Box* over 20%. Figure 2 lists all 12 movies with the number of views higher than 10 million. Although the two top lists are not identical, all movies on these lists made at least 100 million revenue worldwide and hence were blockbuster movies defined by Internet Movie Database (IMDB).

Figure 1. Movies with *Top Box* over 20%

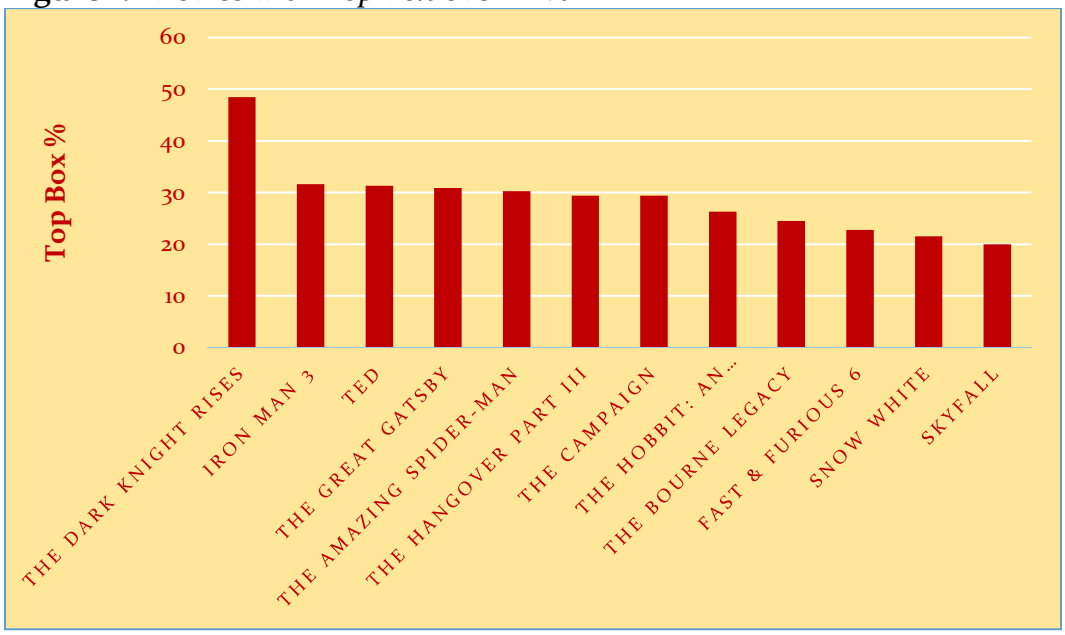
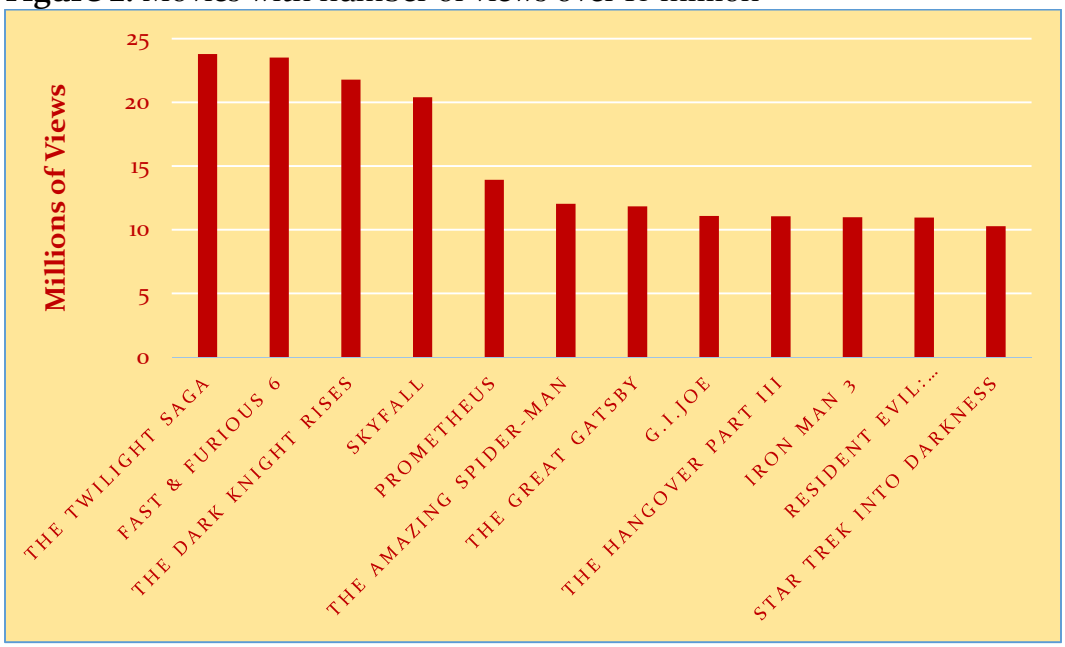


Figure 2. Movies with number of views over 10 million



Prediction model

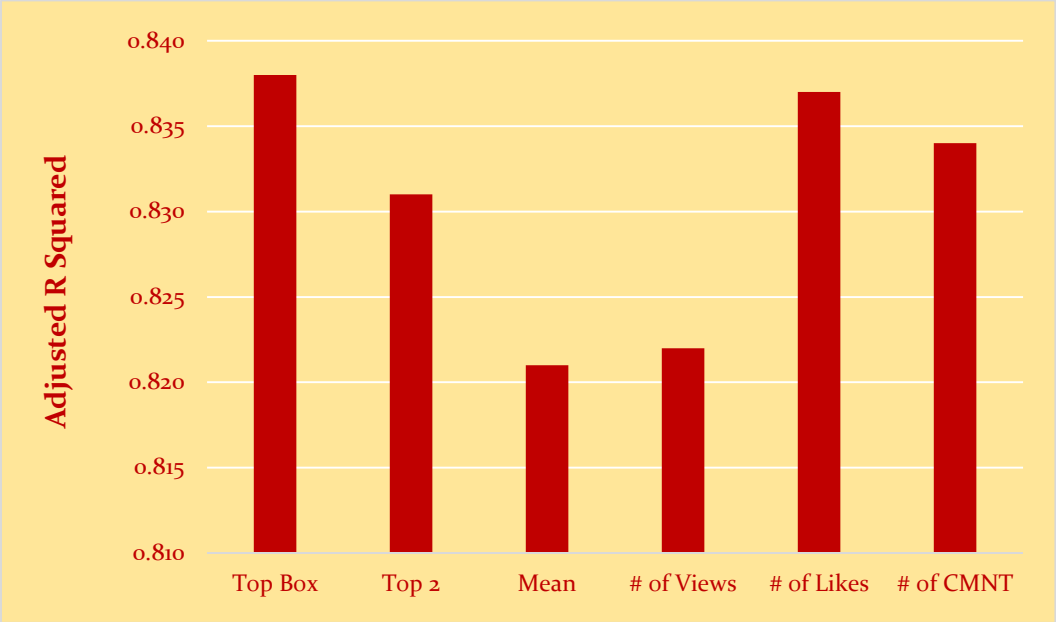
We estimate a log-linear model with the dependent variable being the first week revenue. The independent variables include the following movie factors: budget, number of opening screens, whether the movie is a sequel, the number of stars, whether the movie is released during a holiday weekend

or in the summer, as well as the number of competing movies released in the same week.

Valid predictors

All models have an adjusted R^2 above .82, indicating good fit (see Figure 3). All metrics are positive predictors of movie revenue except the *mean* metric. This is an expected result. In the purchase intention setting, *Top Box* and *Top 2 Box* are the most commonly used intention metrics. Although the mean metric is more popular than the *Top Box* and *Top 2 Box* in other settings (such as a customer satisfaction survey), it is not commonly used in the purchase intention setting. At the individual level, various psychological biases or situational factors can lead to discrepancies between purchase intentions and actual behavior. For example, a change in financial status or even a simple change in mood can affect a respondent's behavior over time. Therefore, only people who show the greatest purchase intention are counted as potential buyers. It is common to make similar adjustments in other settings as well. For instance, for consumer products, it is customary to use a .8 multiplier for the most interested group and .3 for the second most interested group (on a commonly used 5-item scale) while disregarding all other responses. As the model with the *Top Box* measure explains the most variance, further discussion uses *Top Box* as representative of the purchase intention metrics.

Figure 3. Model fit statistics for all metrics.



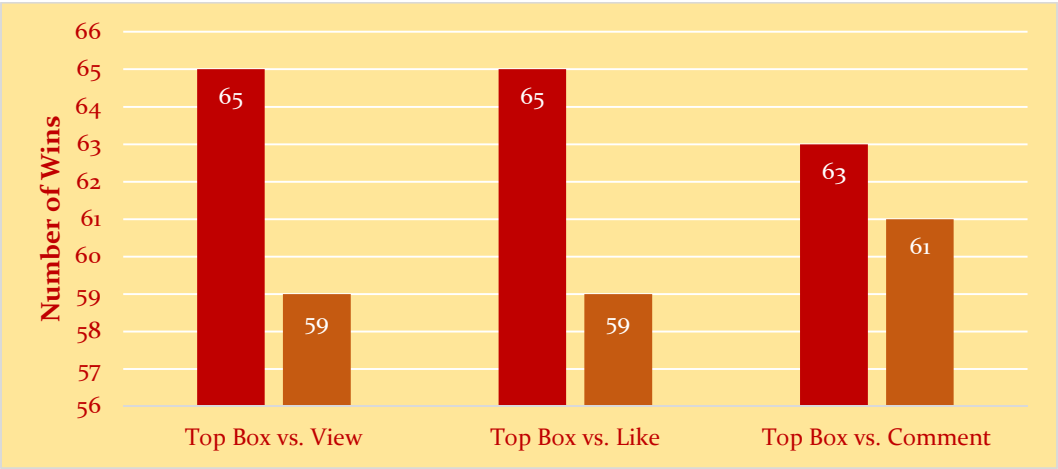
Prediction accuracy comparison

When comparing the prediction accuracy, we count the number of wins of each metric. The number of wins is related to the prediction error of each metric for each movie. A metric is considered to "win" for a specific movie if it has the lower absolute percentage error (APE) compared with the alternative metric. This is a common practice to evaluate the relative accuracy of two approaches.[22]

Why survey approach is better in the movie setting?

Based on the number of wins (see Figure 4), *Top Box* is the best metric. This finding suggests that direct purchase intention elicitation of the survey approach and the ability to differentiate between levels of interest outweigh the sample size advantage of customer buzz. The reason could be that much of the information in the buzz is not relevant in the prediction. After all, the bar for generating buzz is much lower than that for an actual purchase. It takes only several minutes for an individual to view a movie trailer or write some comments on YouTube while spending several hours in a movie theatre represents a much bigger commitment.

Figure 4. The relative accuracy of different metrics based on the number of wins



Another possible reason is that using metrics at the aggregation level reduces the severity of the intention-behavior discrepancy issue attributed to the survey approach. Although the discrepancies are common, they occur at the individual level. Prediction at aggregate level is different from an examination of whether the intention matches the behavior at the individual level. As long as the information contained in the sample represents the characteristics of the general movie-going population, the intention-

behavior relationship at the aggregate level can be expected to be stronger than at the individual level.

In addition, predictions based on the intention survey for new products tend to be less accurate than for existing products because consumers are generally more familiar with existing products, however this finding mostly applies to tangible products.[23] The situation is different for experiential products such as movies, plays, books, and recorded music. Experiential products usually contain some elements that are familiar to customers. For example, each movie contains certain elements that are familiar to the audience such as the director, the screenwriter, the actors, and the major theme.

University students are commonly used in consumer behavior studies by academia and are generally assumed to be representative of the targeted population for most products (e.g., computer, books, monitors, food etc.). However, people might still argue that students are not perfect representation of the population. The superior performance of our study indicates that the representativeness issue may not be serious.

Specialized movie products

As different types of movie are expected to receive different level of customer buzz, it is natural to ask: what is the performance of buzz when a movie receives little buzz? Would the predictive power of buzz remain the same? Hence, we estimated additional regression analyses to identify which factors contribute to the generalizations of customer buzz and purchase intention. In these models, the dependent variables are pre-release buzz metrics or purchase intention metrics, and the independent variables are movie factors as described previously. We find that budget is the only significant predictor across all models. It has a positive impact on both buzz and purchase intention. High budgets are also positively associated with movie revenues.

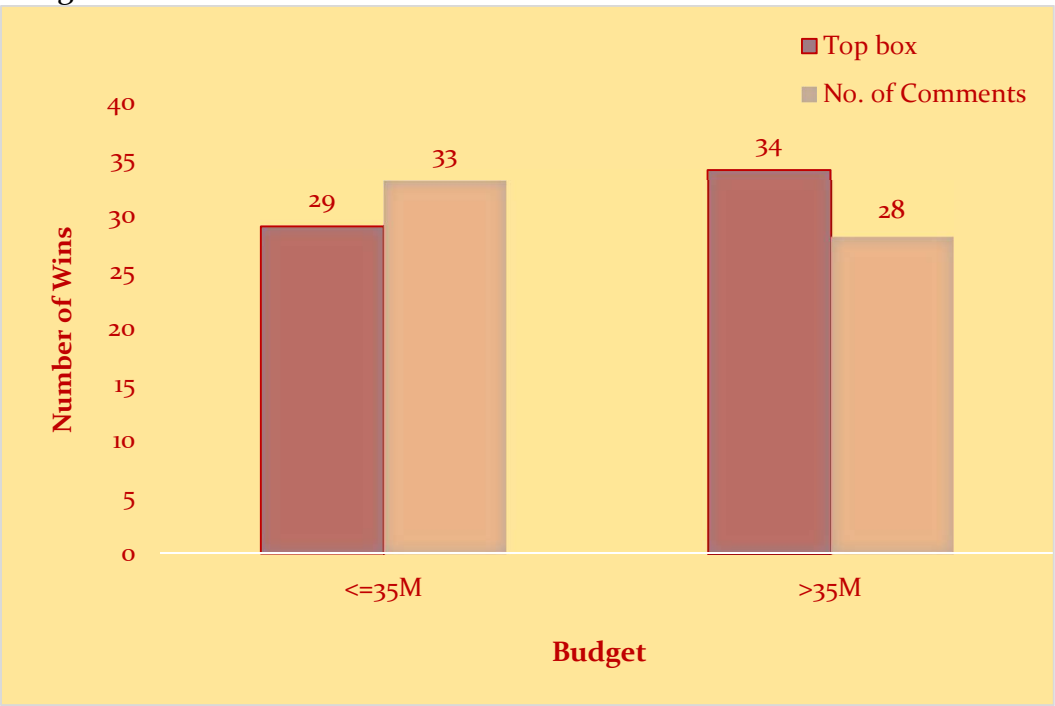
Therefore, we categorize movies based on different levels of budget. Although there is no clear-cut criterion about budget level, a budget in the range of $20 to $50 million is generally considered moderate. Movie industry is considered risky because of the high production and marketing costs involved and the commensurate uncertainty of return. A $20 million budget is modest for a comedy/drama movie but is low for an action or a science fiction movie. For example, *Crouching Tiger, Hidden Dragon* (2000) with its budget of $17 million, grossed $214 million worldwide, and is considered a famous example of a very successful low-budget movie.

The mean and median budget in our movie data are $55 million and $35 million, respectively. We use the median split to categorize movies into big and small categories, i.e., above and below median budget respectively.

Buzz is better in low budget movie categories

Figure 5 compares *Top Box* and buzz metrics in the two budget categories. Results indicate that *Top Box* still performs better than the *comment* metric in big-budget movies (34 vs. 28) based on the number of wins. But in the low budget category, the *comment* metric beats the *Top Box* (33 vs. 29).

Figure 5. The relative accuracy of *Top Box* vs. *Comment* metric across budget categories



This result for the low budget movies is different from the previously reported result for the aggregate sample. For the low-budget movies, both intention metrics and buzz metrics perform worse than for high-budget moves in terms of model fit and prediction error (results are available upon request). This is understandable because these movies are generally more difficult to predict. For survey approach, identifying the right survey respondents for these movies is particularly difficult. Budgets reflect the ambition of the producers. Most low-budget movies are either independent movies or movies made by inexperienced or unknown filmmakers. With more varied "risky" concepts (i.e., higher percentage of novel/unique

concepts) contained in these movies, the acceptance level of the potential target market is more uncertain.

For buzz, these movies usually are not well recognized by the market and hence receive less attention in social media. In our movie data sets, the average number of views for big budget movies is 6.8 times greater than that for small-budget movies. However, for the low budget movies, YouTube achieves a better performance, which can be attributed to the self-selection mechanism. With numerous videos uploaded to YouTube daily, the selection of online videos to watch and comment on has become a daily decision among consumers. The decision itself reflects an individual's own and independent preference towards a product. Only viewers interested in a particular niche genre are likely to view and comment on highly specialized movie trailers on YouTube. The self-selection (i.e., segmentation) mechanism makes YouTube an ideal tool for predicting the performance of highly specialized niche products (e.g., Japanese anime films). This finding highlights a unique advantage of social media platforms: allowing the participants to self-select themselves into a targeted sample, which is otherwise very hard to achieve through other approaches.

Complements or substitutes?

We further investigate whether the two approaches are substitutes or complements. We include both the *comment* metric and *Top Box* metric in a model (the best metrics of each approach). The model with the two metrics (hereinafter referred to as the combined model) performs better than any model with a single metric in terms of the number of wins (Figure 6). The combined model also has the best model fit across all models (adjusted R^2 = .851).[24] Although previously *Top Box* is the best single metric, the combined model outperforms the *Top Box* prediction by more than 20 movies.

Moreover, the combined model has the lowest average predictor error (Figure 7). Combing survey approach and the customer buzz can increase the accuracy by at least 5% in terms of prediction accuracy. Given that "[s]mall forecasting errors can mean a big hit on profits"[25] when it comes new product forecasting, this represents a managerially significant improvement. Considering an average opening week movie revenue of 30 million in year 2012, a 5% prediction error would amount to at least 1.5 million in revenues.

In addition, the combination does not make any of the metrics an insignificant predictor, suggesting that intention and buzz complement each other in improving prediction accuracy. This interesting finding suggests that combining both traditional approach and new approach could substantially improve forecasting accuracy.

Figure 6. Combining metrics vs single metrics

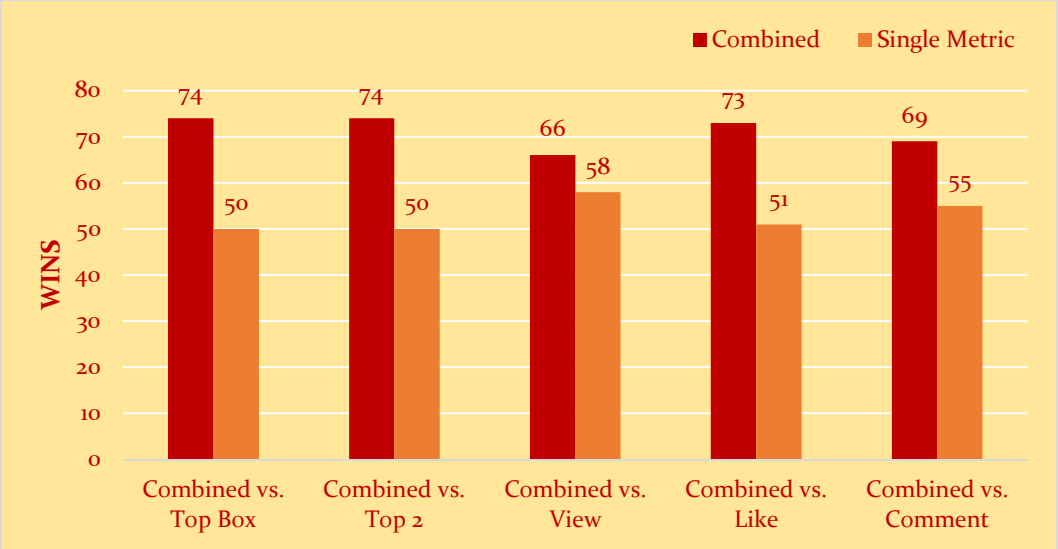
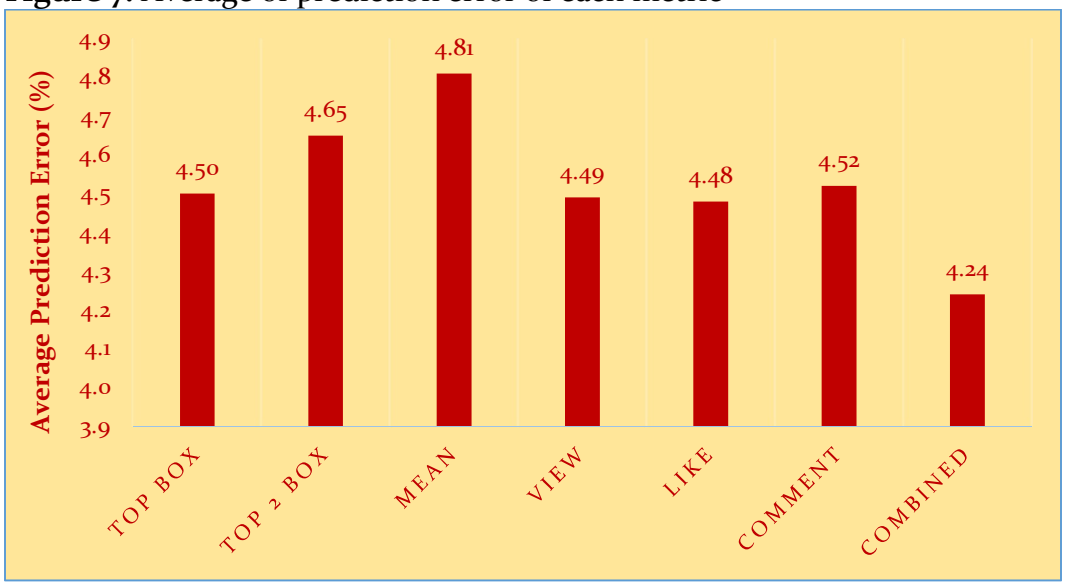


Figure 7. Average of prediction error of each metric



Discussion and Conclusion

Our study provides evidence regarding the relative accuracy of the two approaches. Surveys are an effective tool in movie prediction both at the aggregate level and in the popular product category. Our conjecture is that surveys should be valued more highly within the durable goods category or existing goods category because these products can be viewed as popular products given their familiarity to the general population.

Meanwhile, customer buzz is particularly useful in predicting performance of low budget movies. The contrast between big-budget and

small-budget movies is similar to that between mainstream/popular products and less-popular products. Many new niche products may be unfamiliar to the general population and fall into the less-popular product category where their newness adds additional challenges for prediction. As such, new concepts/features valued by designers may not be well-received by the market. It is worth noting that many new products are considered to be too novel and/or complex for the mass market in the introduction stage (e.g., Macintosh) and may benefit from entrepreneurial marketing to attract attention.[26] Furthermore, many niche products become market hits as they improve user friendliness, remove externalities, and 'go mainstream.' As the success of businesses in competitive environments increasingly rest on their ability to be entrepreneurial and innovative,[27] we are especially interested in the prediction accuracy for these niche products. This finding provides a direction on utilizing buzz to predict the sales performance of new products under conditions of high uncertainty.

Several theories advanced by Ernest Dichter in his seminal work related to WOM communication motives provide explanations for the superior performance of buzz in predicting the sales of less popular products.[28] According to his *self-involvement* or *enhancement* theory, people are motivated to express their opinion about a product due to a need to enhance themselves in front of the audience. This suggests that consumers are more likely to express their opinions when their opinions can "make them look unique, intelligent and savvy."[29] Hence the motivation to express oneself is higher when a product is less popular or even considered an "underdog." Similarly, according to his *other involvement* theory, customers are more likely to express themselves when they have a genuine desire to help others make a good purchase decision. This motivation implies that a customer has to demonstrate a greater knowledge about a product than their audience. It provides the rationale that when a customer is more familiar with an unpopular product than a popular product, he/she is more likely to engage in online discussions.

The results show that social media buzz complements traditional approaches well and it even outperforms and may serve as a substitute for the traditional survey method under conditions of high uncertainty. Overall, the best prediction accuracy is observed when the two approaches are used in tandem, however.

Use of social media data would also be valuable for marketing research classes where our finding can be introduced/demonstrated to the students through an online buzz analysis assignment.[30] We urge marketing researchers to apply the straightforward comparisons described here to their settings and examine the extent to which these findings apply to their

product categories. It may often be possible to significantly improve the prediction outcomes derived from simplistic measures such as the Net Promoter Score[31] by combining it with buzz-based scores.

Authors

Yuying Shi is an assistant professor in the Department of Marketing and Business Analytics at the Texas A & M university-commerce. She received both her Ph.D. in Marketing and Master of statistics from University of Florida. Her research area focuses on marketing analytics and digital marketing. Her research has been published in Journal of the Academy of Marketing Science, Review of Marketing Science, British Journal of Mathematical and Statistical Psychology, etc.
email: yuying.shi@tamuc.edu

Dr. Kate Karniouchina is the Dean of Lorry I. Lokey School of Business and Public Policy. Kate holds a PhD in Marketing, an MBA, and a BA degree in Finance from the University of Utah. Her work has been widely published in academic and industry journals including the Journal of Marketing, Strategic Management Journal, International Journal of Research in Marketing, Journal of Product Innovation Management, Cornell Hospitality Quarterly, Marketing Letters, Journal of Service Management, and European Journal of Operational Research. She is a marketing research expert who carries out projects for a number of small business, corporate and government clients.
email: kkarniouchina@mills.edu

Can Uslay is an Associate Professor and the Vice Dean for Academic Programs and Innovations at Rutgers Business School, Newark and New Brunswick. He received his MBA and Ph.D. from the Georgia Institute of Technology. His research interests lie broadly within marketing strategy and theory construction. He is a recipient of the Chancellor's Award, the Valerie Scudder Award, and several Dean's awards for outstanding scholarship, teaching, and service. His research has been presented in various international conferences and published in the leading academic journals such as the Journal of Marketing, Journal of the Academy of Marketing Science, European Business Review, International Business Review, International Journal of Technology Management, International Journal of Business Environment, International Journal of Quality & Reliability Management, Journal of Advertising Education, Journal of Business-to-Business Marketing, Journal of Business Research, Journal of Business Strategy, Journal of Public Policy & Marketing, Journal of Research in Marketing & Entrepreneurship, Journal of Strategic Marketing, Marketing Education Review, Review of Marketing Research, and the Rutgers Business Review.
email: can.uslay@business.rutgers.edu

Endnotes

1. Schmidt, J., & Calantone, R. (2002). Escalation of commitment during new product development. *Journal of Academy of Marketing Science, 30*(2), 103-118.
2. Hu, K., Acimovic, J., Erize, F., Thomas, D., & Van Mieghem, J.A. (2017, March 10). How to predict demand for your new product. *Kellogg Insight.*
3. Fishbein, M., & Ajzen, I. (1975). Belief, attitude, intention, and behavior. Reading, MA: Addison–Wesley.
4. The percentage of respondents choosing the highest categories.
5. Houston, M.B., Kupfer, A.K., Hennig-Thurau, T., & Spann, M. (2018). Pre-release consumer buzz. *Journal of the Academy of Marketing Science, 46*(2), 338-360.
6. Karniouchina, E.V. (2011). Are virtual stock markets efficient predictors of new product success? The case of the Hollywood stock exchange. *Journal of Product Innovation Management, 28*(4), 470-484.
7. Bond, P. (2014, June 5). Adobe predicts summer 2014's box-office losers. *Hollywood Reporter*.
8. Gibbs, S. (2017, September 22). IPhone 8: Muted reaction and small queues lead to questions over demand. *The Guardian.*
9. IDC forecasts revenues for big data and business analytics solutions will reach $189.1 billion this year with double-digit annual growth through 2022. (2019, April 4). IDC.
10. Social media analytics market by component, application (sales and marketing management, customer experience management, and competitive intelligence), deployment model, organization size, industry vertical, and region - global forecast to 2024. (2019). *Markets and Markets.*
11. Boyd, J. (2020). The complete social listening guide. *BrandWatch.*
12. Engel, B., Blackwell, R., & Kollat, D.T. (1978). *Consumer Behavior.* New York: Rinehart.
13. De Vany, A.S., & Walls, W.D. (2004). Motion picture profit, the stable Paretian hypothesis, and the curse of the superstar. *Journal of Economic Dynamics and Control, 28*(6), 1035-1057.
14. Foutz, N.Z., & Jank, W. (2010). Research note—prerelease demand forecasting for motion pictures using functional shape analysis of virtual stock markets. *Marketing Science, 29*(3), 568-579.
15. Karniouchina, E.V. (2011). Impact of star and movie buzz on motion picture distribution and box office revenue. *International Journal of Research in Marketing, 28*(1), 62-74.
16. Karniouchina, E.V., Uslay, C., & Erenburg, G. (2016). The case for product placement. *Rutgers Business Review, 1*(1), 77-83.
17. Karniouchina, E.V., Uslay, C., & Erenburg, G. (2011). Do marketing media have life cycles? The case of product placement in movies. *Journal of Marketing, 75*(3), 27-48.
18. Chung, C., Niu, S.C., & Sriskandarajah, C. (2012). A sales forecast model for short-life-cycle products: New releases at Blockbuster. *Production and Operations Management, 21*(5), 851-873.
19. Ho, T.H., & Chen, K.Y. (2007). New product blockbusters: The magic and science of prediction markets. *California Management Review, 50*(1), 144-158.
20. Schuman, H., & Presser, S. (1996). *Questions and answers in attitude surveys: Experiments on question form, wording, and context.* Thousand Oaks, C.A.: Sage Publications.
21. Motion Picture Association of America. (2016). *Theatrical Market Statistics Summary 2015*. Motion Picture Association.
22. Spann, M., & Skiera, B. (2003). Internet-based virtual stock markets for business forecasting. *Management Science, 49*(10), 1310-1326.

23. Morwitz, V.G., Steckel, J.H., & Gupta, A. (2007). When do purchase intentions predict sales? *International Journal of Forecasting, 23*(3), 347-364
24. We also tried other combinations of different metrics in each approach (results are available upon request) and obtained consistent results.
25. Hu, K., Acimovic, J., Erize, F., Thomas, D., & Van Mieghem, J.A. (2017, March 10). How to predict demand for your new product. *Kellogg Insight.*
26. Uslay, C. (2017). The good, bad, and ugly side of entrepreneurial marketing: Is your social media campaign unveiled, incognito, or exposed? *Rutgers Business Review, 2*(3), 338-349.
27. Alqahtani, N., & Uslay, C. (in press). Entrepreneurial marketing and firm performance: Synthesis and conceptual development. *Journal of Business Research.*
28. Dichter, E. (1966). How word-of-mouth advertising works. *Harvard Business Review,* 44, 147-166.
29. Dellarocas, C., Gao, G., & Narayan, R. (2010). Are consumers more likely to contribute online reviews for hit or niche products? *Journal of Management Information Systems, 27*(2), 127-158.
30. Malhotra, N.K., Dixit, A., & Uslay, C. (2002). Use of internet technology in marketing research education. *Marketing Education Review, 12*(3), 25-34.
31. Keiningham, T.L., Cooil, B., Andreassen, T.W., & Aksoy, L. (2007). A longitudinal examination of new promoter and firm revenue growth. *Journal of Marketing, 71*(3), 39-51.