Athlete Endorsement Effect in Twitter: Perspective from Big Data

Hyunseok Song, Indiana University
Kevin K. Byon (Advisor), Indiana University

Marketing - Sponsorship (Professional Sport) Thursday, June 7, 2018
Poster 2:10 PM
Abstract 2018-079 Room: Halifax

Corporations have long recognized the importance of athlete endorsement, and as such, significant resources are allocated. Consequently, top athletes earning via endorsement deal more than their salary is not uncommon (Forbes, 2017). Several positive consequences associated with the athlete endorsement have been identified in sport marketing/advertising literature. Notable effects include enhancing consumer's attitude toward the endorsed brand and exerting purchase intentions of the endorsed brand (Cunningham & Bright, 2012; Ruihley, Runyan, & Lear, 2010).

With the advanced technology, social media has become one of the most widely used marketing communication tools as it is related to athlete endorsement (Brison, Baker, & Byon, 2013; McKelvey & Masteralexis, 2013). The new phenomenon is largely attributed to continuously increasing attention. For instance, In 2017, the number of social media users surpassed 2.46 billion (Statista, 2017), and digital ad spending surpassed that of TV (eMarketer, 2016). Of several types of social media, Twitter has garnered a particular attention as an athlete endorsement tool due to its simple characteristic (140 words limitation) and unique function (the hashtag). This growing attention is observed in academic research as a number of scholars have examined the effects of athlete endorsement on Twitter.

Professional athletes’ use of Twitter for the purpose of endorsement (Abeza et al., 2017), legitimacy as an advertising medium (Brison et al., 2013), characteristics of tweets for online word-of-mouth (Cork & Eddy, 2017), the influence of athlete tweets toward consumer’s attitude (Cunningham & Bright, 2012), and athletes’ Twitter hashtag use during a major sporting event (Blaszka, Burch, Frederick, Clavio, & Walsh, 2012). Nonetheless, the extant research suffers from limitations, including but are not limited to the following: (a) use of a cross-sectional data has been a popular trend. limiting our understanding of the true effect of Twitter as the real-time conversation is one distinct attribute of online social media (e.g., Twitter) and (b) employing small sample size, resulting in generalizability issue. Thus, filling these gaps would advance the theoretical knowledge associated with the effect of athlete endorsement via Twitter. The purpose of this study is to examine the athlete endorsement effect via Twitter by adjusting time lag among the variables using longitudinal big data. To collect data, we employed the IUNI Observatory on Social Media (OSoMe). This tool is granted to access approximately 70 billion public tweets from mid-2010 through 2016. The sample size is approximately 10% of total tweets and our samples are assigned randomly. It is built on the Apache Big Data Stack (ABDS) framework with two modules: Hadoop and HBase (Davis et al., 2016). As for typing key words, we could collect time serial number of tweets out of 70 billion tweets.

Data collection procedures are as follows: First, Nike was chosen as the corporate brand because it is one of the largest athlete endorsement companies as the firm currently (at the time of data collection) endorses 47 of the 100 highest paid athletes (Forbes, 2017). Second, the 47 Nike-endorsed athletes were categorized according to six types of sport: baseball, basketball, football, tennis, soccer, and golf. Third, we analyzed the number of hashtags of each key from June 1, 2016 through December 4, 2016 daily. We counted Twitter keywords applying OSome such as #nike, #baseball, #basketball, #football, #tennis, #soccer, and #golf. Additionally, to examine real-time influence of the tweet conversation, we adjusted time lags using the following formulae:

\[
\text{cov}(\text{nike}(t), \text{baseball}(t))/\sqrt{\text{s}(\text{nike}(t))\times\text{s}(\text{baseball}(t)))
\]

\[
\text{cov}(\text{nike}(t), \text{baseball}(t-1))/\sqrt{\text{s}(\text{nike}(t))\times\text{s}(\text{baseball}(t-1)))},
\]

\[
\text{cov}(\text{nike}(t), \text{baseball}(t-2))/\sqrt{\text{s}(\text{nike}(t))\times\text{s}(\text{baseball}(t-2)))},
\]

where t is the date when the tweet was posted, so the tweet at t-1 indicates one day before than the tweet at t. Similarly, the tweet at t-2 means that the tweet is posted two days before than the tweet at t. We examined from zero-lag correlation through two-day-lag correlation between Nike and the types of sports.

The results showed that #baseball and #golf were highly correlated with #nike Corr(nike(t),baseball(t))=0.54, and Corr(nike(t),golf(t))=0.48, and their effect size using the Cohen’s D was large and medium, respectively. However, basketball, football, tennis and soccer were lowly correlated (Corr(nike(t),basketball(t))=0.08,
Corr(nike(t),football(t))=0.09, Corr(nike(t),tennis(t))=0.30, and Corr(nike(t),soccer(t))=0.29). When time lag was adjusted, it was found that correlation values decreased for baseball and golf (Corr(nike(t),baseball(t-1))=0.49, Corr(nike(t),baseball(t-2))=0.47, Corr(nike(t),golf(t-1))=0.26, Corr(nike(t),golf(t-2))=0.31). For basketball, football, tennis and soccer, most correlation values increased slightly on the most, however the values were still low (Corr(nike(t),basketball(t-1))=0.18, Corr(nike(t),basketball(t-2))=0.10, Corr(nike(t),football(t-1))=0.19, Corr(nike(t),football(t-2))=0.22, Corr(nike(t),tennis(t-1))=0.38, Corr(nike(t),tennis(t-2))=0.24, Corr(nike(t),soccer(t-2))=0.33, Corr(nike(t),soccer(t-2))=0.37). The correlation among tweets is stronger at the same time, and time lag makes weaker the correlation. This result supports that online tweet conversation works in real-time, and although, extant literature without considering time lag, they do not need to examine their test more in the time lagged condition.

This study has implications that can contribute to body of knowledge in athlete endorsement effects in social media. First, we used large, objective, and longitudinal-nature data, resulting in more valid results. By adjusting time lag between the endorsed brand (Nike) and the types of sports, we found that time delay actually decreases the magnitude of the endorsement effects via Twitter. This result indicates that Twitter effects may decay as time goes, which can be explicated by decay theory (Berman, 2009). However, we did not know how fast the effects actually fade out as we only adjusted the daily time lag. In the future research, hourly or minutely time lag should be applied. Second contribution is that our results could help marketing managers better understand how their brand and endorsers are congruently perceived as a form of athlete endorsement via Twitter. In this research, baseball and golf are highly correlated with Nike, meaning that tweet users may feel that Nike was more closely align with baseball and golf athlete endorsement. The above implications will be elaborated along with how this line of research using big data can be further developed to advance knowledge in athlete endorsement in the presentation.