Journal of Population Economics (2024) 37:56
https://doi.org/10.1007/s00148-024-01034-7

ORIGINAL PAPER

Do not shut up and do dribble: social media and TV consumption

Matteo Pazzona[1] · **Nicola Spagnolo**[1,2,3]

Received: 12 December 2022 / Accepted: 17 May 2024 / Published online: 30 May 2024

Abstract

This paper investigates the impact of social media interest and sentiment surrounding the 2020 National Basketball Association's involvement with the Black Lives Matter movement on the television audience in the United States. Twitter (now known as X) serves as the chosen social media platform, and we determine the sentiment expressed in tweets (messages posted on Twitter) using the XLM-RoBERTa deep language model. Our primary findings indicate that the quantity of users' posts does not significantly influence TV viewership; instead, the tone of the messages plays a crucial role. Positive messages supporting the NBA's engagement correlate with an increase in the number of viewers, while those expressing opposition do not. We argue that this asymmetry may stem from a positive elasticity among casual (non-habitual) NBA viewers concerning positive sentiments toward NBA involvement. These viewers are likely to align with the NBA's stances on civil rights and BLM. In contrast, the core NBA fan base exhibits inelastic demand and is unlikely to cease watching NBA games. A comprehensive set of robustness checks reinforces the validity of our key conclusions.

Keywords Black lives matter · Sentiment · Social media · TV consumption

JEL Classification D12 · D80 · L82

Responsible editor: Kompal Sinha

✉ Matteo Pazzona
matteo.pazzona@brunel.ac.uk

Nicola Spagnolo
nicola.spagnolo@brunel.ac.uk

1 Department of Economics and Finance, Brunel University London, London, Middlesex UB8 3PH, UK

2 Centre for Applied Macroeconomic Analysis (CAMA), Australian National University, Canberra, Australia

3 Dipartimento di Economia, Universitá degli studi della Campania "Luigi Vanvitelli", Naples, Italy

1 Introduction and motivation

Traditionally, global corporations tend to maintain a low profile on contentious issues like immigration or racism (Smith and Korschun 2018). However, in recent times, companies seem increasingly willing to adopt a more assertive stance, actively participating in social and political discourse. This shift is notably facilitated by the widespread influence of social media platforms, which amplifies the reach and impact of corporate positions.

The media can influence public behavior through various mechanisms, including the salience effect (or exposure effect), where extensive discussions make a topic highly noticeable (McCombs and Shaw 1972; Erbring et al. 1980). Additionally, the frame effect, which concerns the tone of a discussion, plays a role in shaping perceptions (Scheufele 1999). While the impact of traditional media (TV and newspapers) on individual opinions and consumption behavior is well-documented in political science and communication literature (McCombs and Shaw 1972; Erbring et al. 1980; Entman 1993), the influence of social media on consumption has received less attention.[1] This study aims to contribute to this understanding, focusing on the 2020 engagement of the National Basketball Association (NBA) with the Black Lives Matter (BLM) movement. Specifically, we investigate the extent to which public opinion, represented by Twitter social media data, influenced the choices of sports consumers, measured through TV ratings and viewership numbers.

A considerable number of consumers now rely on social media platforms such as Facebook, Instagram, and Twitter for news and information, gradually replacing or supplementing more traditional media outlets like TV and newspapers. Among these platforms, Twitter stands out with over 330 million active monthly users, making it one of the most influential channels and a primary source for real-time information.[2] Twitter allows users to share concise and frequent messages ("tweets"), limited to 280 characters. These messages are easily searchable and can be quickly shared (retweeted) with one's followers. As highlighted by Austmann and Vigne (2021), discussions on Twitter have the potential to reach a broader audience beyond its user base. Twitter data has been extensively utilized in various research domains (Soroka et al. 2018; Shen et al. 2019; Zhuravskaya et al. 2020). Importantly, it is the most commonly used social platform by journalists and professional players in major American sports leagues.

The Black Lives Matter (BLM) movement emerged in 2013 as a social movement protesting police brutality against black individuals in the USA. In June 2020, widespread protests erupted across the country following the death of George Floyd, catalyzing one of the most significant civil rights movements in recent American history. The National Basketball Association (NBA) has been notably vocal and supportive of the BLM cause. Upon the resumption of play in late July 2020, after a COVID-19-induced suspension, the NBA adopted a firm but contentious stance, allowing athletes to kneel during the national anthem, wear social justice messages on their

[1] Social media has played a central role in shaping sports development in this century, with Baimbridge et al. (1996) being one of the first to explore this relationship, examining broadcasting and football using UK data.

[2] As of the time of writing, Twitter has been rebranded as X. However, we consistently refer to it as Twitter, considering the name during the study period.

jerseys instead of surnames, and display the BLM slogan on basketball courts.[3] Many prominent players and NBA Commissioner Adam Silver endorsed the Black Lives Matter protests, with Silver stating on June 1, 2020: “Together with our teams and players, we will continue our efforts to promote inclusion and bridge divides.”

While the NBA’s engagement with the BLM movement aimed to raise awareness of social issues, it faced criticism from some politicians. In September 2020, President Donald Trump tweeted: “People are tired of watching the highly political NBA. Basketball ratings are way down, and they won’t be coming back.” President Trump’s statement implied that consumers may penalize companies for political involvement. NBA Commissioner Adam Silver responded, stating: “No data BLM on-court hurts NBA ratings. There is no doubt there are some people who have become further engaged with the league. They respect their right to speak out on issues that are important to them.” This polarization was amplified by social media users. Motivated by these dynamics, our study aims to answer two key questions: To what extent did individuals alter their NBA TV consumption due to the interest generated by the league’s involvement (intensity) in the BLM movement, and was this change driven by the tone (sentiment) of messages posted on social media?

The empirical nature of our research questions means that we do not hold any preconceived expectations regarding the impact on consumers, if such an impact exists. Individuals might have chosen to: reduce their TV consumption due to disapproval of the NBA’s involvement in non-sport-related issues; increase their time spent watching NBA matches, indicating support for the league’s social stance; or remain unaffected, indicating a low elasticity of NBA TV demand to social media commentary.

To disentangle the role of social media from other concurrent factors, we initially conducted an event study, utilizing high-frequency data and incorporating a comprehensive set of control variables. Our measure of the TV audience included data on viewers and ratings for all NBA matches broadcast on national TV networks during the 2018–2019 and 2019–2020 seasons. While TV audience data is not a traditional measure of consumption, it is closely linked to revenues. Kanazawa and Funk (2001), for instance, found that higher ratings for locally televised NBA basketball games enabled teams to generate greater advertising revenues.

To gauge social media sentiment toward the NBA’s involvement in the BLM movement, we utilized XLM-RoBERTa-base (referred to as XLM-RoBERTa). This is a pre-trained natural language processing technique (Conneau et al. 2019) trained on a large corpus of English data in a self-supervised manner. A key feature of XLM-RoBERTa is its ability to comprehend the meaning of nuanced languages, including the interpretation of emoticons. Using XLM-RoBERTa, we predicted the likelihood of positive, neutral, or negative sentiments in tweets.

We constructed both an intensity (exposure) index, representing the total number of tweets on the NBA’s involvement with the BLM movement, and a sentiment (tone) index, categorizing Twitter data into positive and negative sentiment clusters. In our baseline empirical analysis, we conducted regressions using TV audience data for

[3] It is worth noticing that before the start of the 2020-2021 season, the NBA commissioner announced the discontinuation of social justice messages on both courts and jerseys. We capitalize on the specific time frame (June - November 2020) during which the NBA took this position to assess its potential impact on viewers and ratings.

NBA matches as the dependent variable and included the social media intensity and sentiment indexes as explanatory variables. Our results highlight that only positively toned tweets have a significant impact on sports TV consumption. Notably, our findings suggest an average viewership increase of 6.6% during the period following the resumption of play after the COVID-related hiatus. These results remain robust when using different model functional forms and tweet sentiment classification metrics. Additionally, we exploited the granularity of Twitter data by exploring the popularity of messages and authors.

An interpretation of the asymmetric effects suggests that NBA core viewers, characterized by an inelastic demand, continue to enjoy watching basketball regardless of public opinions. Conversely, the number of marginal (non-habitual) viewers appears to increase with positive sentiment toward the NBA's involvement, indicating an elastic demand. This group is likely to tune in due to their alignment with the social campaigns presented by the NBA, thus supporting the economic significance of our findings.

To account for potentially omitted variables, we supplemented our analysis with a difference-in-differences (DiD) approach, using the National Hockey League (NHL) as a control. The NHL serves as an ideal comparison to the NBA, sharing the following features: both leagues have the same number (82) of regular-season matches; the regular season for both leagues was interrupted and resumed at the same time; and, finally, both leagues exhibited a similar (negative) trend in TV viewers during the period under consideration (2018–2020). In line with our primary findings, there is limited evidence suggesting that the intensity of NBA involvement in the BLM movement contributed to the decline in the number of viewers and ratings. While the nature of our data does not allow inferences about changes in behavior at the individual level, the results confirm that aggregate TV sports demand is inelastic with respect to the total number of tweets referring to the NBA's involvement in the BLM movement.

We conclude our empirical analysis by examining a proxy for traditional media intensity, measured by the number of articles appearing in newspapers, as well as TV and radio transcripts. While the Twitter-based indexes gauge the active involvement of social media users on a certain topic, motivating the focus of this paper, consumers of traditional media are passively exposed to the authors' views. Additionally, we considered the volume of Google searches on the NBA-BLM involvement using daily Google Trends data. Consistent with the previous results, we do not find any statistically significant effect for any of the non-social-media-intensity indexes.

Our work draws inspiration from various strands of the economics, politics, and marketing literature. Firstly, we examine whether media coverage and its tone affect consumer behavior (Carroll 2003; Lerner et al. 2007; Lamla and Lein 2014; Biolsi and Lebedinsky 2021; Abdollahi 2023).[4] Secondly, we explore beliefs and consumption patterns (Angeletos and La'O 2013; Gillitzer and Prasad 2018; Benhabib and Spiegel 2019). These studies acknowledge the influence of non-fundamental factors, such as beliefs and opinions, on consumers' spending decisions. Within this context, recent literature has found that social media is increasingly influencing beliefs and fostering activism among consumers (Bovitz et al. 2002; Hendel et al. 2017; Enikolopov et al.

[4] Similarly, (Ananyev et al. 2021) showed how exposure to Fox News Channels affected physical distancing during COVID-19.

2020; Zhuravskaya et al. 2020; Gorodnichenko et al. 2021). Thirdly, we explore the demand for sports, especially for TV broadcasting (Hausman and Leonard 1997; Forrest et al. 2005; Buraimo et al. 2022; Caselli et al. 2024). This literature focuses on aspects such as outcome uncertainty, team identity, and the quality of players. Very few studies have analyzed the role of other determinants.[5] Finally, this study is also linked to recent marketing literature that examines the influence of social and political activism on a firm's performance (Scherer et al. 2014; Smith and Korschun 2018). We contribute to this by disentangling the involvement of firms from that of other confounding factors.

The paper's structure is organized as follows. Section 2 introduces and examines the data. Section 3 outlines the results of the baseline model. Section 4 explores a series of robustness checks. Section 5 reports the findings using difference-in-differences methodology. Section 6 presents additional results using traditional media outlets. Finally, Section 7 provides concluding remarks.

2 Data description

In this section, we describe the variables employed in the empirical analysis.

2.1 TV audience

Major professional sports leagues in the USA operate as private entities and are not publicly traded on the stock market. Consequently, using share value as an indicator of a firm's success is not applicable in this context. Instead, our approach involves gauging the NBA's performance through two alternative indicators: TV viewership and ratings. While these measures are distinct, they are interconnected and provide insights into the television audience (Nielsen 2021). "Ratings" denote the percentage of U.S. TV households tuned in to a specific program, while "TV share" represents a percentage based on the households engaged in television viewing. On the other hand, "viewers" and "viewing figures" encompass the total number of individuals watching a program. It is important to note that TV ratings and viewing figures may not comprehensively capture all aspects of media consumption, given the evolving landscape of how people consume TV, particularly with the increasing prevalence of programs on internet-related platforms. Unfortunately, data pertaining to this aspect is not available for our analysis.[6]

[5] Buraimo et al. (2016) — the closest to our work — documented the negative effect of an (in)famous scandal — *Calciopoli* — on stadium attendance for the sanctioned teams. However, that scandal had only a negative connotation — as it concerned some clubs colluding with referees — whereas in our setting, involvement in the BLM could be either supported or not supported by fans.

[6] The primary income stream for professional sports leagues is derived from the revenue generated through national TV deals. In 2021, these deals amounted to a substantial $2.7 billion, constituting over a third of the total revenue for the NBA. Additionally, a discernible positive correlation exists between the TV audience and advertising revenues, (Kanazawa and Funk 2001). The revenue composition for professional sports clubs has undergone significant changes (Buraimo and Simmons 2015). In the short term, this information can guide broadcasters in selecting which matches to broadcast and influence companies' decisions regarding the amount to bid for advertising slots during the game (Buraimo et al. 2022).

The TV scheduling of NBA games in the USA aligns with other major professional sports leagues, such as the National Football League and the National Hockey League. A limited number of games are nationally broadcasted, while the remainder are accessible through regional networks. Our analysis centers on the nationally broadcasted games, as detailed data on regional network broadcasting is unavailable. These nationally aired games are typically featured on cable TV channels like ABC and TNT. Consumers typically subscribe to a package, involving payment, that encompasses a bundle of TV channels providing various content such as news and movies. Once subscribed, users enjoy unlimited access to all programs broadcast by those networks. This subscription fee, from an economic standpoint, constitutes a sunk cost with an almost negligible marginal cost of consumption. This setup is advantageous for our empirical analysis because consumers can promptly adjust their viewership behavior without incurring additional costs. The decision to subscribe can be viewed as the extensive margin, while the actual consumption reflects the intensive margin. Our data and empirical setting are better suited for investigating the latter.

We collected data on all 506 NBA matches broadcast by national networks during the 2018–2019 and 2019–2020 seasons. The dataset for ratings includes information on 490 matches, encompassing details such as date and time, broadcasting network, and the game type (regular season or playoff). It is important to note that certain days featured more than one televised match. Table 1 reveals an average viewership of 2.35 million for these matches, contrasting with an average rating of 1.47. Approximately 68% of the matches were aired during prime time (after 8 pm ET), with 33% occurring in the playoffs and 20% on weekends. *ESPN* and *TNT* were the primary broadcasters for the majority of the games.[7]

2.2 Social media intensity and sentiment

To measure the intensity (or exposure) and public sentiment (or tone) regarding the NBA's involvement with BLM, we utilized data from the social media platform Twitter. Our approach involved searching for all original Twitter messages containing both the terms "NBA" and "BLM," encompassing variations such as abbreviations or fully spelled-out forms. Specifically, our search criteria included terms like "BLM," "#BLM," "Black Lives Matter," "Black Live Matters," and so forth.[8] Our focus was solely on original tweets, aligning with the methodology employed in the existing literature (Hatte et al. 2021). All tweet-related information was obtained from the Twitter Research Access API, and we specifically considered messages written in English or those with an undefined language, provided that hashtags were included.[9] Consequently, we created the variable *Tweets NBA∩BLM* to represent the total daily

[7] To save space, the aforementioned figures are omitted from Table 1. Additionally, Table 1 provides essential summary statistics for the NHL, serving as a basis for comparison with the NBA in the difference-in-differences setting (Section 5). Our data-set includes viewing-figure data for 414 NHL matches and ratings data for 278 matches. It is important to note that all TV audience data is aggregated at the national level, as data for lower administrative levels was not accessible.

[8] The outcomes align with stricter criteria, such as incorporating only tweets featuring the NBA and BLM hashtags, as illustrated in the Appendix.

[9] More details are available at https://developer.twitter.com/en/use-cases/do-research/academic-research.

Table 1 Summary statistics

	Observations	Mean	SD	Min	Max
NBA					
Viewers (millions)	506	2.35	2.14	0.25	18.76
Ratings	490	1.47	1.23	0.17	10.70
Tweets NBA∩ BLM(-1d)	506	51.04	165.17	0	1436
Tweets NBA∩ BLM(-1d) Positive	506	12.14	57.59	0	581
Tweets NBA∩ BLM(-1d) Negative	506	13.33	37.50	0	296
Tweets NBA∩ BLM(-1d) Neg&Neu	506	38.56	112.18	0	853
Tweets NBA(-1d)	506	4011.46	1474.05	1267	10738
Tweets BLM(-1d)	506	892.49	1499.32	17	8101
Tweets(-1d) Positive Afinn	506	22.67	84.54	0	820
Tweets(-1d) Negative Afinn	506	18.30	56.68	0	490
NYT&USA&Fox(-1d)	506	0.26	0.75	0	7
News Transcripts(-1d)	506	1.25	4.40	0	44
Google Trends(-1d)	506	15.09	27.72	0	100
Tweets NBA∩ BLM(-1d), Pop	506	1942.76	8824.26	0	69465
Tweets NBA∩ BLM Pos(-1d), Pop	506	606.19	3103.01	0	28953.90
Tweets NBA∩ BLM Neg(-1d), Pop	506	456.09	1937.01	0	16845.32
Tweets NBA∩ BLM(-1d), Pop (Followers, xK)	506	7067.98	44400	0	4.78E+05
Tweets NBA∩ BLM Pos(-1d), Pop (Followers, xK)	506	2503.04	18210.01	0	198808.30
Tweets NBA∩ BLM Neg(-1d), Pop (Followers, xK)	506	1490.24	8097.28	0	84202.87
Tweets NBA∩ BLM, Leag+Pla(-1d)	506	0.40	1.11	0	7
Tweets NBA∩ BLM, Fans(-1d)	506	50.63	164.48	0	1432
After	506	0.24	0.43	0	1
Play-off	506	0.33	0.47	0	1
Prime Time	506	0.68	0.47	0	1
Weekend	506	0.20	0.40	0	1
NHL					
Viewers (millions)	414	0.72	0.85	0.12	0.87
Ratings	278	0.56	0.54	0.14	4.90
Tweets NHL∩ BLM(-1d)	414	19.75	65.48	0	349

Notes: This table provides summary statistics for the key variables employed in this study count of original tweets, serving as our indicator for intensity or exposure. Illustrated in the upper panel of Fig. 1, the tweet count remained consistently minimal until the end of May 2020, after which it experienced a sharp and notable surge. Two distinct peaks were discernible: the first occurred following the resumption of NBA matches at the end of July 2020, while the second was associated with the boycott following the police shooting of a black man in Wisconsin in August of the same year.

The variable *Tweets NBA∩BLM* measures the intensity of public interest regarding the NBA's engagement with BLM. This intensity could influence TV viewership through two contrasting channels. On one hand, some viewers might opt not to watch matches due to disagreement with the NBA's involvement. On the other hand, viewers might increase their viewership to express support. In essence, a potential explanatory factor for sports consumption behavior lies not solely in the quantity of tweets but rather in their tone. This is particularly relevant in the current social media era where polarization on sensitive issues has become commonplace (Dawson et al. 2014; Lamla and Lein 2014). Consequently, we conducted a sentiment analysis on each message within the *Tweets NBA∩BLM* variable.

A conventional objective in sentiment analysis involves determining whether the tone conveyed in a given text is positive, negative, or neutral. The advent of advanced language models like BERT and RoBERTa enables the exploration of more intricate data domains, such as texts where authors tend to express their opinions or sentiments less explicitly, or even rely on the use of emoticons (Ho et al. 2020; Hamborg et al 2021). In this study, we employ the XLM-RoBERTa language model, developed by Facebook AI (Conneau et al. 2019). This model integrates cross-lingual language models (XLMs) with RoBERTa, where RoBERTa stands for Robustly Optimized BERT

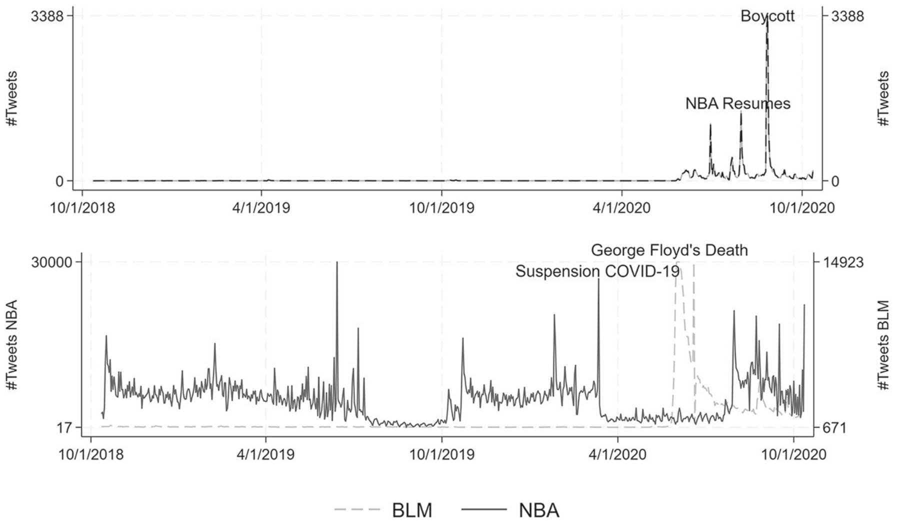

Fig. 1 Trends in tweets. Notes: The top panel reports the number of *Tweets NBA∩BLM*, whereas the bottom panel reports the number of *Tweets NBA* and *Tweets BLM*

Approach (Liu et al. 2019).[10] Much like BERT, RoBERTa employs a transformer-based architecture and undergoes pre-training on extensive textual data, enabling it to acquire contextual representations of words and phrases. However, in contrast to BERT, RoBERTa is trained on a more extensive data-set and follows a more efficient training procedure.[11] The NLP model utilized in this study, XLM-RoBERTa, undergoes pre-training on 2.5TB of filtered CommonCrawl data encompassing 100 languages. For tweet classification, we employed the general-purpose Python library TweetNLP (Camacho-Collados et al. 2022). TweetNLP relies on the transformer-based language models RoBERTa and XLM-R as its backbone, further pre-trained on Twitter-specific corpora. In addition to sentiment analysis, TweetNLP supports various tasks, including emotion recognition, irony detection, etc.[12] The model employed in TweetNLP assigns a positive, negative, or neutral probability to each tweet, with these probabilities summing up to one.

We categorize a tweet as positive, negative, or neutral based on the sentiment with the highest probability. Consequently, we construct the variables *Tweets Positive*, *Tweets Negative*, and *Tweets Neutral* by summing the daily tweets belonging to each sentiment group.[13] The top panel of Fig. 2 presents the histogram of *Tweets NBA*$\cap$*BLM* from June 1, 2020, categorized by sentiment type. In the bottom panel, we illustrate the average sentiment probability over time. Throughout the entire period, neutral sentiment maintains a slightly higher probability. Negative and positive sentiments exhibit comparable probabilities, indicating a high level of polarization with neither sentiment dominating the other.

We also collected tweets containing the word NBA but not BLM, and those containing BLM but not NBA. Two variables, labeled *Tweets NBA* and *Tweets BLM*, were constructed to represent the daily total number of tweets. These measures capture public interest in the NBA and BLM separately. In the bottom panel of Fig. 1, we depict the trends of *Tweets NBA* and *Tweets BLM*. The former follows a cyclical path, with more tweets posted at the start of the regular season and during the playoffs. In contrast, *Tweets BLM* peaked immediately after the death of George Floyd. A comparison between the bottom and top panels reveals that the NBA's involvement with BLM did not perfectly align with the trends of *Tweets NBA* and *Tweets BLM*. This observation is reassuring, suggesting that NBA/BLM-related tweets follow a distinct trend.

3 Baseline model results

Our baseline model has the following form:

$$TV\,Consumption_{i,t} = \alpha + \beta TweetsNBA \cap BLM_{i,t-1} + \gamma X_{i,t} + \varepsilon_{i,t}, \qquad (1)$$

[10] BERT (Bidirectional Encoder Representations from Transformers) is a distinct language model developed by Devlin et al. (2018).

[11] BERT and RoBERTa undergo pre-training on raw texts without human labeling.

[12] To accomplish this, additional models such as TweetEval, TimeLMs, and XLM-T were utilized (Barbieri et al. 2020, 2021; Loureiro et al. 2022).

[13] In Table 4, alternative approaches to aggregate XLM-RoBERTa-related probability sentiments are explored.

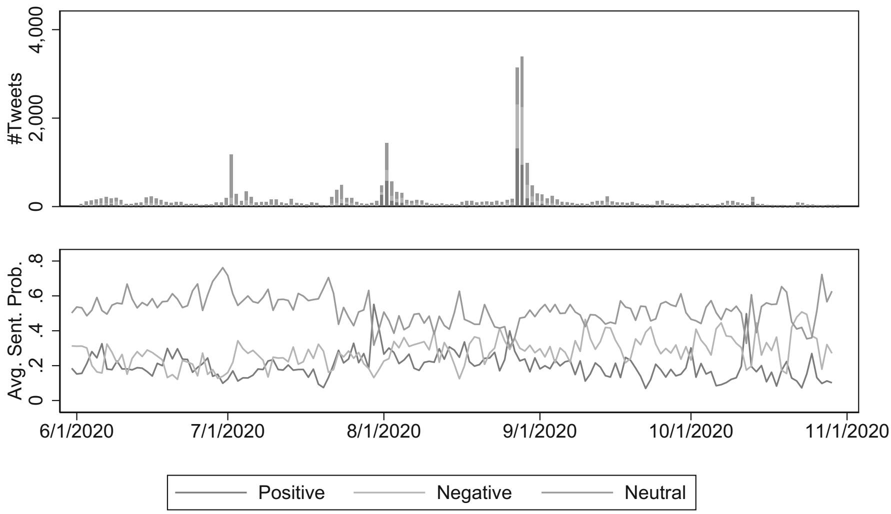

Fig. 2 Tweets and average sentiment probability. Notes: This figure illustrates the number of tweets posted from June 1, 2020, onwards. Top panel: histogram of the total number of *Tweets NBA∩BLM* divided by the type of sentiment. Bottom panel: average daily probability associated with each sentiment - positive, negative, and neutral

where i represents the match number, and t captures the daily time dimension. The variable *TV Consumption* denotes the number of viewers (in logs) and ratings, taken in turns.[14] *Tweets NBA∩BLM* reflects the NBA/BLM intensity, serving as a proxy for social media interest in the topic. This variable has been scaled by 1,000 to enhance the interpretation of the coefficients. We incorporate the lagged value, denoted as $t-1$, under the assumption that information requires time to disseminate and reach social media users. Individuals typically need some time to process and subsequently respond to social media inputs. Additionally, due to the nature of our data, we were unable to precisely control for whether tweets were released on the same day before or after NBA matches were aired.[15]

The variable set X encompasses a comprehensive array of control variables pertinent to both intensity and TV audience. *Tweets NBA* and *Tweets BLM* were detailed in the previous section. *Key Dates* is a binary variable with a value of one on days marked by significant events correlated with media exposure and audience impact. One such event is the restart of NBA matches in July 2020, during which players took a knee for the first time. Another key date is associated with the August 2020 boycott, where several NBA teams refused to play following the shooting of Jacob Blake in Wisconsin. Additionally, we incorporate two dummy variables capturing critical comments made by President Trump regarding the NBA's engagement with the BLM movement. One corresponds to September 1, when he tweeted the message outlined

[14] We did so to account for the severe right skewness of the viewership variable. Results with viewers in levels are qualitatively similar and are not reported but available upon request.

[15] The use of lagged variables also helps mitigate potential issues related to reverse causality, as highlighted by Reed (2015). Nevertheless, we also estimated our baseline model using tweets posted on the same day, and the unreported results showed qualitative consistency.

in the introduction. The second dummy is linked to the well-known interview given to the TV cable channel Fox on August 5.[16] To account for the COVID-19 context, we introduce the dummy variable *After*, taking a value of one for matches played after the resumption of the NBA at the end of July 2020 (refer to Fig. 1, top panel). The inclusion of this dummy aims in interpreting the role of *Tweets NBA∩BLM* during the COVID-19 period.

Using regular season matches as the reference group, we incorporate controls for different play-off stages (1st and 2nd rounds, semi-finals, and finals). Dummy variables are employed to account for matches aired during prime time and on weekends. A similar approach is applied to the three primary TV networks broadcasting the highest number of matches: *ESPN*, *ABC*, and *TNT*. Additionally, we introduce a linear daily trend and season-fixed effects into the model. Residuals are clustered, and *p*-values are computed using wild bootstrap (Cameron and Miller 2015).

The outcomes are displayed in Table 2, columns (1) and (4), corresponding to *Viewers* and *Ratings*, respectively. The point estimates do not show statistical significance at conventional levels, indicating that social media exposure did not impact *TV Consumption*. Following the examination of the total number of tweets, our focus now shifts to their tone (or sentiment). We consider the subsequent specification:

$$
\begin{aligned}
TVConsumption_{i,t} = \alpha + \zeta TweetsPositive_{i,t-1} + \delta TweetsNegative_{i,t-1} \\
+ \gamma X_{i,t} + \varepsilon_{i,t}, \qquad (2)
\end{aligned}
$$

where $Tweets\ Positive$ and $Tweets\ Negative$ are the variables detailed in Section 2. The findings are presented in columns (2) and (5) of Table 2. In columns (3) and (6), we replace negative tweets with $Tweets\ Negative\&Neutral$, representing the sum of negative and neutral tweets. A discernible pattern emerges, suggesting that the tone of the message plays a crucial role. Days featuring a higher number of positively toned tweets are associated with an increase in both *Viewers* and *Ratings*, while negatively toned tweets are linked to a decrease. However, only the results for positive tweets attain statistical significance, whereas those for negative tweets do not. Quantifying the effect, the log-linear model in Column (2) indicates that for the average number of positively toned tweets, the increase in the number of viewers is 1.16%. However, for games played post the COVID-related hiatus (a more meaningful time frame) the increase in viewership amounts to 6.6%. Alternatively, if we focus only on the three days with the highest number of tweets, the effect is comparable to that of the initial round of playoffs. We deem this effect economically significant as it captures the behavior of casual (non-habitual or marginal) fans who mobilize to watch NBA matches in support of its involvement in the BLM movement.[17] Our results also align with the perspective that responses to positive and negative messages are asymmetrical

[16] https://www.foxnews.com/video/6178381393001

[17] To interpret the coefficient's magnitude, exponentiation is necessary, along with applying the formula $(exp(\zeta) - 1) * 100\%$. This formula provides an approximation of the percentage change in *Viewers* for a one-unit increase in the regressor *Tweets (-1d) Positive*. Specifically, an increase of 1,000 positively toned tweets corresponds to a 258% increase in *Viewers*. To calculate the effect for values other than 1,000, we use the formula $[(exp(c\zeta) - 1) * 100\%]$, where *c* represents the fraction of 1,000 tweets under consideration.

Table 2 Social media intensity, sentiment and TV consumption: baseline model

	Viewers(Log)			Ratings		
	Intensity (1)	Only Negative (2)	Negative Neutral (3)	Intensity (4)	Only Negative (5)	Negative Neutral (6)
Tweets NBA∩ BLM(-1d)	0.33			0.62		
	(1.84)			(1.96)		
Tweets(-1d) Positive		1.28*	1.38*		2.73*	3.19**
		(2.21)	(2.68)		(2.41)	(3.06)
Tweets(-1d) Negative		−1.05			−2.63	
		(−0.78)			(−1.00)	
Tweets(-1d) Neg&Neut			−0.37			−1.10
			(−1.11)			(−1.55)
Tweets BLM(-1d)	0.01	0.06	0.05	0.09	0.20	0.19**
	(0.44)	(1.16)	(1.57)	(1.18)	(1.62)	(2.02)
Tweets NBA(-1d)	−0.01	−0.01	−0.01	−0.01	−0.02	−0.02
	(−0.45)	(−0.56)	(−0.60)	(−0.24)	(−0.43)	(−0.49)
After	−0.13	−0.22	−0.19	−0.90***	−1.10***	−1.06***
	(−0.77)	(−1.21)	(−1.22)	(−2.54)	(−2.66)	(−2.88)
Play-Off:1st Rd	0.55***	0.56***	0.56***	0.69***	0.72***	0.74***
	(8.30)	(8.62)	(8.43)	(5.82)	(5.94)	(5.97)
Play-Off:Semi	1.11***	1.12***	1.12***	1.58***	1.58***	1.59***
	(15.51)	(15.20)	(15.49)	(11.80)	(11.69)	(11.93)
Play-Off:Conf Finals	1.49***	1.48***	1.49***	2.40***	2.40***	2.40***
	(20.60)	(19.99)	(20.26)	(14.29)	(13.94)	(13.96)
Play-Off:Finals	1.53***	1.55***	1.55***	4.64***	4.69***	4.69***
	(12.01)	(12.40)	(12.26)	(7.25)	(7.39)	(7.36)
ABC	1.74**	1.74***	1.74***	1.30*	1.30	1.31
	(4.08)	(4.09)	(4.09)	(2.01)	(2.01)	(2.03)
ESPN	1.06*	1.06	1.06	0.21	0.22	0.23
	(2.13)	(2.14)	(2.14)	(0.29)	(0.30)	(0.31)
TNT	1.02	1.03	1.03	0.21	0.21	0.22
	(2.06)	(2.07)	(2.07)	(0.29)	(0.30)	(0.31)
Prime Time	0.23***	0.23***	0.23***	0.21***	0.21***	0.21***
	(6.13)	(6.13)	(6.15)	(4.00)	(4.01)	(4.04)
Weekend	0.03	0.04	0.04	−0.18	−0.17	−0.17
	(0.52)	(0.63)	(0.64)	(−1.40)	(−1.33)	(−1.27)
Key Dates	−0.06	−0.08	−0.07	−0.10	−0.15	−0.13
	(−0.44)	(−0.63)	(−0.57)	(−0.39)	(−0.61)	(−0.51)
Season FE, Trend	Yes	Yes	Yes	Yes	Yes	Yes

Table 2 continued

	Viewers(Log)			Ratings		
	Intensity (1)	Only Negative (2)	Negative Neutral (3)	Intensity (4)	Only Negative (5)	Negative Neutral (6)
Observations	506	506	506	490	490	490
Adj. R-sq	0.66	0.66	0.66	0.72	0.73	0.73

Notes: All social media intensity measures are based on the day preceding the matches (-1d). *Tweets NBA∩BLM* reflects the count of original tweets obtained by using "NBA" and "BLM" as keywords, including their spelled-out forms. *Tweets Positive* and *Tweets Negative* pertain to the subset of *Tweets NBA∩BLM* with tones in favor or against the NBA's engagement with the BLM movement, classified using the XLM-RoBERTa deep language model. *Tweets NBA* represents the number of tweets posted the day before the matches using the hashtag "NBA" but not "BLM." *Tweets BLM* considers tweets with the hashtag "BLM" but not "NBA." Descriptions of other variables can be found in the text. Residuals are clustered at the league level, and *p*-values are calculated through wild bootstrap with t-statistics reported in parentheses. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

(Soroka 2006; Akhtar et al. 2011). To highlight our key findings, we visually present the main coefficients in Fig. 3.

The remaining variables, namely *Tweets NBA* and *Tweets BLM*, consistently lack statistical significance. The variable *After* reveals a decline in TV consumption upon the NBA's resumption of play, but this effect is only significant for *Ratings*. Matches during the playoff period attracted a larger audience, with the Finals registering the

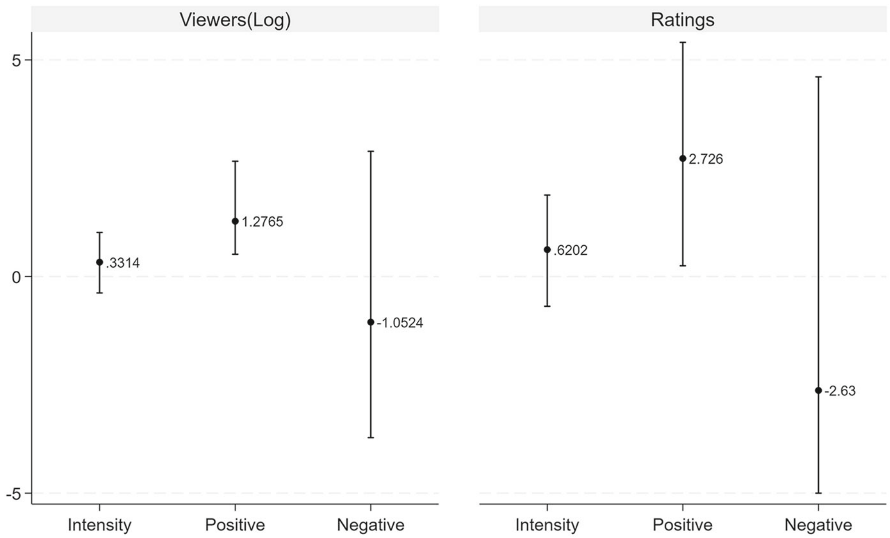

Fig. 3 Main results. Notes: This figure illustrates the coefficients for *Tweets NBA∩BLM*, *Tweets Positive*, and *Tweets Negative* as reported in Table 2's columns (1), (2), (4), and (5). The confidence intervals are based on wild bootstrapping and are presented asymmetrically. The lower-bound confidence interval associated with *Ratings* is truncated at -5

highest value. As anticipated, the variable *Prime Time* exhibits a positive and significant impact.

By incorporating a comprehensive set of variables and using high-frequency data, along with lagged values, we aim to provide a robust identification strategy. It remains plausible to contend that the intensity of tweets related to NBA/BLM is endogenous concerning the TV audience. For instance, a match involving the Chicago Bulls and the Los Angeles Lakers might attract more interest - ceteris paribus - than a less prestigious one. Similarly, games played between teams in urban, multicultural areas might draw different audiences compared to those in more rural, homogeneous areas. To account for such characteristics, in Table 3, we replicate the baseline results by incorporating match-up fixed effects.[18]

We present initial results for *Viewers* followed by those for *Ratings*. Our findings indicate that the overall impact of social media intensity on *TV Consumption* is negligible. It is the sentiment of the tweets that holds substantial importance, where tweets supporting the NBA's engagement with BLM are linked to a larger audience. However, the incorporation of match-up fixed effects results in an absolute increase in the coefficients capturing the relationship between tweet sentiment and the number of viewers. The statistical significance is also heightened. For instance, a significant effect is observed for *Tweets(-1d) Neg&Neu*.

4 Robustness analyses

In this section, we provide a series of exercises that are intended to challenge the results presented in the baseline analyses.

4.1 Alternative metrics of sentiment

The initial robustness-check exercise explores alternative approaches to weighing tweet sentiments. In the baseline model, tweets are categorized as positive, negative, or neutral based on their respective probabilities. The sum of all tweets within each sentiment category is then calculated for each day. However, in the baseline model, a tweet might be labeled as positive even if the negative sentiment has only a slightly lower probability. To address this, we introduce the variables *Tweets NBA∩BLM Positive, Tweets NBA∩BLM Negative*, and *Tweets NBA∩BLM Neutral*. These variables are created by multiplying the average daily probability for each sentiment group by the total number of tweets posted in a day. Consistent with our baseline specification, we explore two models: one including only negative tweets and another combining neutral and negative tweets.

In addition, we utilize another widely-used lexicon called *Afinn*, which assigns sentiment scores to 2,476 English words on a scale from -5 (indicating the most negative sentiment) to +5 (indicating the most positive sentiment). Tweets with scores

[18] The term "match-up" refers to a game played more than once by the same two teams. In our data, we have a total of 248 match-ups, which results in a substantial reduction in degrees of freedom and explains our decision not to consider them in the baseline model.

Table 3 Social media intensity, sentiment, and TV consumption: including game FE

	Viewers(Log)			Ratings		
	Intensity (1)	Only Negative (2)	Negative Neutral (3)	Intensity (4)	Only Negative (5)	Negative Neutral (6)
Tweets NBA∩ BLM(-1d)	0.41			0.89		
	(1.48)			(1.79)		
Tweets(-1d) Positive		1.68*	2.55**		3.51**	5.45***
		(1.93)	(2.56)		(2.32)	(3.21)
Tweets(-1d) Negative		−0.49			−0.59	
		(−0.26)			(−0.19)	
Tweets(-1d) Neg&Neut			−0.83			−1.71*
			(−1.39)			(−1.68)
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes
Match-Ups FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	506	506	506	490	490	490
Adj. R-sq	0.75	0.75	0.75	0.77	0.77	0.77

Notes: This table reports the baseline results while controlling for match-up fixed effects. All measures of media intensity pertain to the day before the matches were played (-1d). *Tweets NBA∩BLM* indicates the count of original tweets obtained using the “NBA” and “BLM” keywords, including their spelled-out forms. *Tweets Positive* and *Tweets Negative* signify subsets of *Tweets NBA∩BLM* with tones in favor of or against the NBA’s engagement with the BLM movement, respectively. The tone was classified using the XLM-RoBERTa deep language model. *Tweets NBA* represents the number of tweets posted on the day before the match with the hashtag “NBA” but not “BLM.” *Tweets BLM* considers tweets with the hashtag “BLM” but not “NBA.” Descriptions of other variables are provided in the text. Residuals are clustered at the league level, p-values are calculated using wild bootstrap, and t-stats are presented in parentheses. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

between -5 and 0 are classified as negative, while those with scores between 1 and 5 are considered positive. We compute a metric representing the daily count of tweets associated with each specific sentiment.[19] The results from these analyses, presented in Table 4, align with those in Table 2. The magnitude of the point estimates indicates that messages with a positive tone had a more significant impact on the TV audience.

4.2 Social media echo chamber

We will now further exploit the granularity of social media data. The primary regressor in Table 2 — *Tweets NBA∩BLM* — represents the quantity of original tweets. However, this measure does not necessarily convey: a) the reception of these tweets by Twitter users; or b) the influence of prominent authors. To address the first concern, in columns (1) to (4) of Table 5 for *Viewers* — and (5) to (9) for *Ratings* — we introduce alternative metrics based on the popularity of the authors of the messages.

[19] To derive this metric, we first calculate the average daily score of the tweets and then multiply it by the total number of tweets, denoted as *Tweets NBA∩BLM*. Our findings remain consistent even when using smaller bins, such as considering tweets scoring -5, -4, and -3 as negative, and +3, +4, and +5 as positive.

Table 4 Alternative methodologies to classify tweet sentiments

	Viewers(Log)			Ratings		
	Negative (1)	BERT Neg&Neu (2)	BERT Afinn (3)	Negative (4)	BERT Neg&Neu (5)	BERT Afinn (6)
Tweets(-1d) Positive	1.41*	1.55*	0.87*	3.33**	4.22**	2.46*
	(2.35)	(2.67)	(2.31)	(2.74)	(3.17)	(2.53)
Tweets(-1d) Negative	−1.50		−0.74	−4.27		−3.17
	(−1.09)		(−0.94)	(−1.51)		(−1.52)
Tweets(-1d) Neg&Neut		−0.54			−1.95*	
		(−1.40)			(−2.11)	
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	506	506	506	490	490	490
Adj. R-sq	0.66	0.66	0.66	0.73	0.73	0.73

Notes: In columns (1), (2), (4), and (5), tweets are categorized as positive, negative, or neutral based on the XLM-RoBERTa deep language model. *Tweets Neg & Neu* represents the combined count of negative and neutral tweets. Columns (3) and (6) display results utilizing the *Afinn* lexicons. All models incorporate *Tweets NBA*, *Tweets BLM*, *Key Dates*, *After*, binary variables for the playoff series, network dummies, *Prime Time*, *Weekend*, *Key Dates*, trends, and season effects. Residuals are clustered at the league and date level, p-values are computed using wild bootstrap. The t-stats are presented in parentheses. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

In columns (1) and (5), we introduce *Tweets NBA∩BLM Pop*, which aggregates tweets, retweets, and likes. This variable is 38 times larger than *Tweets NBA∩BLM*, with a notably high standard deviation.[20] In columns (2) and (6), we further break down *Tweets NBA∩BLM Pop* into positive and negative tweets. Similar to the baseline model, we observe that positively toned tweets are linked to an increase in both *Viewers* and *Ratings*.

To account for the popularity of the authors of the messages, in columns (3) and (7), we multiply the daily intensity of original tweets by the average number of Twitter followers of the authors posting on that day. In columns (4) and (9), we separate this overall popularity measure into positive and negative messages. The results indicate that none of these metrics exhibits a statistically significant effect on *TV Consumption*.

Lastly, in columns (5) and (10), we explore the influence of the authors' identity in the tweets. We introduce the variables *Tweets NBA∩BLM, Leag+ Pla* and *Tweets NBA∩ BLM, Fans*. The former signifies the total count of tweets from accounts associated with NBA stakeholders, such as the league, NBA teams, and NBA players. This metric serves as a close approximation to the NBA's direct support for the BLM movement, reflecting the unfiltered endorsement of the league. On the other hand, *Tweets NBA∩BLM, Fans* represents tweets generated by fans/individuals who are not stakeholders.

This analysis uncovers a positive and statistically significant impact of *Tweets NBA∩BLM, Leag+ Pla* on both *Viewers* and *Ratings*, whereas no discernible effect is observed for *Tweets NBA∩BLM, Fans*. If anything, the active engagement of NBA

[20] The correlation between the two metrics is 0.75.

Table 5 Social media eco chamber

	Viewers(Log)					Ratings				
	Tweets Retweets and Likes (1)	Tweets Retweets and Likes (2)	Tweets Weighted by Followers (3)	Tweets Weighted by Followers (4)	Players + League vs. Fans (5)	Tweets Retweets and Likes (6)	Tweets Retweets and Likes (7)	Tweets Weighted by Followers (8)	Tweets Weighted by Followers (9)	Players + League vs. Fans (10)
Tweets NBA∩ BLM(-1d), Pop	0.00		0.01			0.00		0.02		
	(1.09)		(2.06)			(0.80)		(2.74)		
Tweets(-1d), Pop Pos		0.03*		0.00			0.05*		0.00	
		(2.97)		(1.52)			(3.88)		(0.09)	
Tweets(-1d), Pop Neg		−0.02*		−0.00			−0.05		0.00	
		(−3.00)		(−0.76)			(−4.04)		(0.78)	
Tweets NBA BLM, Leag+Pla(-1d)					0.52*					1.41**
					(2.53)					(3.83)
Tweets NBA BLM, Fans(-1d)					0.28					0.60
					(1.78)					(2.10)
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	506	506	506	506	506	506	506	506	506	506
Adj. R-sq	0.65	0.66	0.66	0.66	0.66	0.72	0.72	0.72	0.72	0.72

Notes: In columns (1) and (6), *Tweets NBA∩BLM, Pop* represents the total number of original tweets along with the related retweets and likes. *Tweets NBA∩BLM Pos, Pop* and *Tweets NBA∩BLM Neg, Pop* in columns (1) and (6) are subsets of *Tweets NBA∩BLM, Pop* scaled by sentiment using XLM-RoBERTa. Similarly, in columns (3), (4), (8), and (9), *Tweets NBA∩BLM* is multiplied by the average number of followers of the Twitter authors who posted on that day. *Tweets NBA∩BLM, Leag+Pla* and *Tweets NBA∩BLM, Fans* represent the number of tweets posted by NBA stakeholders and NBA fans. *Tweets NBA* denotes the number of tweets the day before the match using the hashtag NBA excluding BLM. *Tweets BLM* includes tweets with hashtags BLM excluding NBA. All models include *Key Dates*, *After*, binary variables for the playoff series, network dummies, *Prime Time*, *Weekend*, *Key Dates*, trends, and season effects. Residuals are clustered at the league level, p-values are calculated using wild bootstrap, and the t-stats are displayed in parentheses. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

stakeholders played a mobilizing role in attracting viewers, irrespective of the intensity of that engagement.[21]

5 Difference-in-differences

The findings presented thus far indicate that days marked by heightened public interest in the NBA's involvement with BLM did not result in significant changes in media consumption. To further account for potentially omitted or time-varying factors, we employ a difference-in-differences (DiD) model, using the National Hockey Association (NHL) as a comparative reference. The NHL serves as an optimal counterfactual for the NBA due to several shared similarities. Firstly, both leagues operate from October to June in non-pandemic years. Secondly, the NBA and NHL both suspended play in March 2020 and resumed in July 2020, as illustrated in the top panel of Fig. 2. Thirdly, both leagues witnessed a decline in TV viewership in the preceding seasons. Despite these resemblances, the two leagues significantly differed in their engagement with the BLM movement, as evidenced by the summary statistics in Table 1.[22] We are aware that other major leagues might be slightly more similar to the NBA in terms of demographic - including race - and political affiliation, compared to the NHL.[23] The National Football League was not considered, given the limited number of matches played and the absence of overlap in the regular season.

Figure 4 illustrates the temporal dynamics of *Viewers* for both seasons and leagues. The initial subset encompasses the entire 2018–2019 season, the second focuses on the early part of the 2019–2020 season up to the suspension of activities due to the COVID-19 pandemic; the last subset covers the period from the end of July to October 2020, when the seasons resumed. This figure illustrates that the average number of viewers was higher for the NBA compared to the NHL.[24] However, both leagues exhibit a similar pre-event trend, supporting the application of a difference-in-differences approach. To validate the visual presentation, indicating a shared trend, we estimate the following model:

$$TV\ \ Consumption_{l,i,t} = \alpha\, Tweets\ League \cap BLM_{l,i,t} + \theta_l + \gamma X_{l,i,t} + \varepsilon_{l,i,t}, \quad (3)$$

, where l represents the league, denoted as either NBA or NHL. *Tweets NHL∩BLM* refers to tweets containing the keywords NHL and BLM. The other variables have been discussed in Section 3. Additionally, we incorporate a control variable, *Tweets NHL*,

21 This result may also be associated with the literature on the influence of role models on individual behavior (Farina and Pathania 2020).

22 The NHL did not issue any explicit statements about the movement and rarely addressed racial issues directly. This is noteworthy despite some players expressing support for the BLM movement on their personal social media platforms. Additionally, players decided to boycott four playoff games in November following the shooting of Jacob Blake. Similar to the NBA, however, two games in the 2020 Stanley Cup finals were postponed on August 27 and 28.

23 There are four major professional leagues in USA/Canada: the NFL, NHL, MLB, and NBA.

24 The United States experienced varying restrictions on social distancing and public life, differing by state, county, and city. Peak restrictions were in place in late March and early April 2020, preceding the NBA's engagement with BLM, as shown in Fig. 1

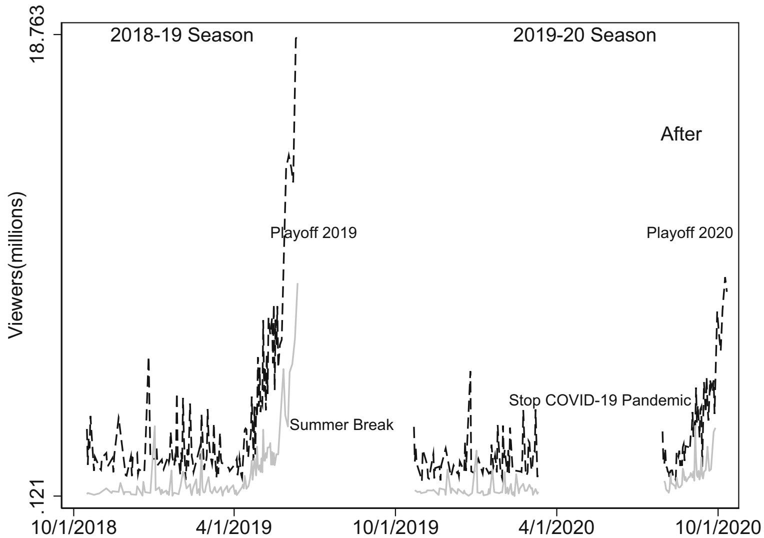

Fig. 4 Trends in viewers and tweets for the NBA and NHL. Notes: This figure provides a visual representation of the temporal dynamics of *Viewers* (in millions) for both seasons and leagues, NBA (black) and NHL (grey). The first sub-sample encompasses the entire 2018–2019 season; the second illustrates the earlier part of the 2019–2020 season until the end of the suspension of all activities due to the COVID-19-related halt (30 July 2020); the last sub-sample (*After*) covers the period from the end of July to October 2020 (i.e., when both seasons resumed)

measuring the number of tweets with NHL excluding those with BLM. The outcomes are presented in columns (1) and (2) of Table 6. The residuals are clustered at the league and date levels, and *p*-values are determined using the wild bootstrap method (Cameron and Miller 2015). The findings underscore that while θ is positive, it is not statistically significant for viewership, whereas it attains significance for *Ratings*.

In columns (3) and (4), we run a standard difference-in-differences model employing a dummy variable that captures the interaction between *NBA* and *After*, labeled as *Int*. This approach accommodates the possibility that the NBA's engagement with BLM leads to a change in our dependent variable over the entire season.[25] Our findings indicate that the NBA did not undergo a decline in viewership (or experience reduced ratings) as a consequence of its involvement with the BLM movement.

6 Conventional media & 2SLS

In this section, we explore intensity measures derived from alternative media platforms, beginning with the use of traditional news outlets (Baker et al. 2016; Caporale et al. 2022). We include one left-leaning outlet, one centrist outlet, and one right-leaning

[25] It is important to note that the two analyses presented in Table 6 are conceptually distinct. In the one utilizing *Tweets NBA∩BLM* and *Tweets League∩BLM*, the treatment variable reflects social media intensity. In the analysis with the dummy variables (columns 3 and 4), we assume a one-off effect throughout the later period. The former aligns with the appropriate robustness exercise for our baseline analysis.

Table 6 Difference-in-differences

	Levels		Dummy	
	Viewers(Log)	Ratings	Viewers(Log)	Ratings
	(1)	(2)	(3)	(4)
Tweets NBA∩ BLM(-1d)	0.02	0.82***		
	(0.27)	(10.48)		
NBA X After			−0.29	−0.14
			(−2.39)	(−1.22)
NBA	0.32	0.85	0.39	0.87
	(5.33)	(9.88)	(5.25)	(26.83)
After	−0.22***	−0.38	−0.04***	−0.39***
	(−3.35)	(−0.88)	(−5.56)	(−1.40)
Other Controls	Yes	Yes	Yes	Yes
Observations	920	768	920	768
Adjusted R-squared	0.77	0.66	0.77	0.66

Notes: Columns (1) and (2) consider the number of tweets related to the involvement of the NBA (NHL) with the BLM movement. Columns (3) and (4) represent the standard DiD model with a dummy controlling for the treated group, *NBA*. *After* is a dummy taking the value one after the NBA and NHL resumed, in and after late July 2020. All models include controls used in the baseline model. Residuals are clustered at the league level, p-values are calculated using wild bootstrap, and t-statistics are reported in round brackets. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

outlet on the US political spectrum.[26] Specifically, for the left-leaning outlet, we utilize *The New York Times*, *USA Today* for the center, and *Fox News* for the right.[27] Similar to the Twitter data, we conducted searches for daily articles or transcripts containing the keywords "NBA" and "BLM," including their spelled-out forms, using Nexis, a news and business research database. The variable *NYT&USA&Fox* was created to encompass all three sources together. The correlation coefficient between the latter and *Tweets NBA∩BLM* stands at 0.65.

Additionally, we collected data on the broadcast transcripts from various US national television and radio news programs, obtained through Nexis. This dataset encompasses networks such as ABC, CBS, CNN, and others. Given the inclusion of *Fox News* in the previous measure, we opted not to incorporate it in this context.[28] We denoted this variable as *News Transcripts*, and it displays a correlation of 0.78 with *Tweets NBA∩BLM*.

The final metric of public interest intensity that we examine is *Google Trends*, developed by Google. This index has found applications across various research domains, including IT, communications, medicine, health, business, and economics

[26] We relied on the media bias ratings provided by the website AllSides, accessible at https://www.allsides.com/unbiased-balanced-news.

[27] Although Fox News is not a newspaper, there are few widely popular conservative-leaning newspapers available.

[28] Results including *Fox News* exhibit qualitative similarity and can be provided upon request.

Table 7 Conventional media & 2SLS

	Viewers(Log)			Ratings			2SLS	
	Mainstream Newspaper (1)	TV & Radio (2)	Web Search (3)	Mainstream Newspaper (4)	TV & Radio (5)	Web Search (6)	Second Stage (7)	First Stage (8)
NYT&USA&Fox(-1d)	−0.02			−0.01				
	(−0.53)			(−0.27)				
News Transcript(-1d)		−0.00			−0.02			0.02
		(−0.33)			(−1.20)			
Google Trends(-1d)			0.00			0.02*		
			(0.83)			(1.69)		
Tweets NBA BLM(-1d)							−0.11	
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	506	506	506	490	490	490	506	506
C-D F stat								147.49
Adjusted R-squared	0.65	0.65	0.65	0.72	0.72	0.73	0.65	0.80

Notes: Results employing three conventional media-intensity measures are presented. *NYT&USA& Fox* is the total number of articles appearing in *The New York Times*, *USA Today*, and *Fox News* websites, using the keywords "NBA" and "BLM." *News Transcripts* refer to the number of these keywords appearing in national TV and radio network transcripts. This variable excludes *Fox News*. Conventional media newspaper articles, along with TV and radio transcripts, were retrieved from the *Nexis* website. *Google Trends* is the daily value of Google intensity searches. In columns (7) and (8), we show the first and second-stage results of the baseline model using *News Transcripts* as an instrument. C-D F stat is the Cragg-Donald Wald F statistic. Residuals are clustered at the league level, *p*-values are calculated using wild bootstrap, and t-stats are reported in parentheses. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

(Vosen and Schmidt 2011; Choi and Varian 2012; Jun et al. 2018).[29] We gather data on internet searches related to the topic "NBA-BLM." The correlation coefficient between *Google Trends* and *Tweets NBA∩BLM* stands at 0.61.

The outcomes for the three indicators of conventional media exposure are presented in Table 7. No compelling evidence of a significant statistical impact on *TV Consumption* is identified. The coefficient for *NYT&USA&Fox* is negative, reflecting the prevailing negative sentiment in articles and transcripts.[30] Notably, only *Google Trends* displays a positive and statistically significant association with *Ratings*.

To further validate our findings, we employ a 2SLS regression model following a methodology employed by Chalfin and McCrary (2018) in examining the impact of police on crime in the USA. These researchers demonstrated that when two variables represent noisy measures attempting to capture the same phenomenon, it is feasible to use one as an instrument for the other. In our study, we utilize alternative measures of media exposure as instruments for *Tweets NBA∩BLM*. Initially, we regress the latter on all three alternative media measures: *NYT&USA&Fox*, *News Transcripts*, and *Google Trends*. Although this analysis is not displayed, it indicates that *News Transcripts* is the most substantial predictor of *Tweets NBA∩BLM*. Consequently, we conduct a just-identified 2SLS model, employing *News Transcripts* as an instrument, while retaining the use of *NYT&USA&Fox* and *Google Trends* as a robustness check (reported in the Appendix). The results, presented in columns (7) and (8), show the absence of an effect on *TV Consumption*. The instrument exhibits a high correlation with *Tweets NBA∩BLM*, and the F-statistic surpasses the threshold suggested by Stock et al. (2002).[31]

7 Conclusions

This paper investigates how the intensity and sentiment of social media posts about the 2020 National Basketball Association's involvement with the Black Lives Matter movement affected the TV audience. Utilizing data from the widely used social media platform Twitter as a measure of media exposure, we find that the posting intensity is not linked to viewership and ratings. However, positively toned posts are correlated with both TV audience metrics, and the observed effect is economically significant. Conversely, the impact of negatively toned tweets is not statistically significant. Our results remain robust across different metrics for constructing social media sentiment proxies and alternative model specifications.

[29] Google Trends offers a scaled measure of the search volume for a particular topic over a defined time period. Specifically, it generates a random sample of searches during the specified duration. Subsequently, it calculates the ratio of topic searches for each day to the total number of Google searches on that day. This ratio is then scaled to 100 for the day with the highest ratio and 0 for the day with insufficient topic searches. Therefore, Google Trends provides the intensity of searches rather than the absolute number.

[30] We utilized the XLM-RoBERTa language model to classify the sentiments of the variable *NYT&USA&Fox* However, the overwhelming majority of articles and news transcripts were negative. This led us to forgo reporting this analysis, as its results would have closely mirrored those in Table 7.

[31] It is important to note that, given the absence of a testable theoretical model, our results should be interpreted as correlational rather than causal effects.

These findings carry significant implications. Firstly, they underscore the relevance of social media in either supporting or opposing the NBA's involvement in social causes, revealing an asymmetry in sentiments. Our sentiment analysis highlights the polarization of opinions on this subject, indicating that the NBA's engagement was not costly, at least in the short term. Secondly, our analysis provides additional insights into the elasticity of demand for sports concerning non-economic factors. We argue that this asymmetry may stem from a positive elasticity among casual (non-habitual) NBA viewers concerning positive sentiments toward NBA involvement. These viewers are likely to align with the NBA's stances on civil rights and BLM. In contrast, the core NBA fan base exhibits inelastic demand and is unlikely to cease watching NBA matches.

We acknowledge the limitation of our study due to the use of national-level data, which might conceal potential heterogeneous patterns at the state and individual characteristics levels. Future research could explore the implications of social media sentiment on sports viewership, considering individual characteristics and preferences. A data-set collected through surveys could be well-suited for this purpose.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00148-024-01034-7.

Acknowledgements The title of this paper is inspired by the phrase "Shut Up and Dribble," used by journalist Laura Ingraham to criticize NBA player LeBron James for discussing politics. We express our gratitude to editor Kompal Sinha and two anonymous referees for their valuable comments and suggestions. Additionally, we extend thanks to Jan Auerbach, Marco Barassi, Olivia Cassero, Alejandro Corvalan, Peter Dawson, Roman Gauriot, Giambattista Rossi, Juan de Dios Tena, and Jan Van Ours for their suggestions. Special acknowledgment goes to John Lewis, the founder of Sports Media Watch, for providing the TV audience data. We appreciate the assistance of Angelo Azzu, Pablo Bustamente, Fuzzy Kara-Isitt, and Exequiel Padilla for their support with Python and R codes. Finally, our thanks go to Federico Buffa, Flavio Tranquillo, and Dan Peterson for making the NBA our favorite sport and for inspiring us to explore the social and economic aspects of the American professional basketball league.

Data availability Data will be made available upon request.

Declarations

Conflict of interest The authors declare no competing interests.

References

Abdollahi H (2023) Oil price volatility and new evidence from news and Twitter. Energy Econ 122:106711

Akhtar S, Faff R, Oliver B, Subrahmanyam A (2011) The power of bad: the negativity bias in Australian consumer sentiment announcements on stock returns. J Banking & Finance 35:1239–1249
Ananyev M, Poyker M, Tian Y (2021) The safest time to fly: pandemic response in the era of Fox News. J Population Econ 34:775–802
Angeletos G-M, La'O J (2013) Sentiments. Econometrica 81:739–779
Austmann LM, Vigne SA (2021) Does environmental awareness fuel the electric vehicle market? A Twitter keyword analysis. Energy Econ 101:105337
Baimbridge M, Cameron S, Dawson P (1996) Satellite television and the demand for football: a whole new ball game? Scottish J Political Econ 43:317–333
Baker SR, Bloom N, Davis SJ (2016) Measuring economic policy uncertainty. Quarterly J Econ 131:1593–1636
Barbieri F, Anke LE, Camacho-Collados J (2021) Xlm-t: multilingual language models in twitter for sentiment analysis and beyond. arXiv:2104.12250
Barbieri F, Camacho-Collados J, Neves L, Espinosa-Anke L (2020) Tweeteval: unified benchmark and comparative evaluation for tweet classification. arXiv:2010.12421
Benhabib J, Spiegel MM (2019) Sentiments and economic activity: evidence from US states. Econ J 129:715–733
Biolsi C, Lebedinsky A (2021) Can changes in sentiments influence consumer behavior? Evidence from the Trump-Russia investigation. Econ Inquiry 59:1569–1592
Bovitz GL, Druckman JN, Lupia A (2002) When can a news organization lead public opinion?-Ideology versus market forces in decisions to make news. Public Choice 113:127–155
Buraimo B, Forrest D, McHale IG, Tena JdD (2022) Armchair fans: modelling audience size for televised football matches. European J Operational Res 298:644–655
Buraimo B, Migali G, Simmons R (2016) An analysis of consumer response to corruption: Italy's Calciopoli scandal. Oxford Bullet Econ Stat 78:22–41
Buraimo B, Simmons R (2015) Uncertainty of outcome or star quality? Television audience demand for English Premier League football. Int J Econ Business 22:449–469
Camacho-Collados J, Rezaee K, Riahi T, Ushio A, Loureiro D, Antypas D, Boisson J, Espinosa-Anke L, Liu F, Martínez-Cámara E, (2022) Tweetnlp: cutting-edge natural language processing for social media. arXiv:2206.14774
Cameron AC, Miller DL (2015) A practitioners guide to cluster-robust inference. J Hum Resourc 50:317–372
Caporale GM, Ali FM, Spagnolo F, Spagnolo N (2022) Cross-border portfolio flows and news media coverage. J Int Money Finance 126102638
Carroll CD (2003) Macroeconomic expectations of households and professional forecasters. Quarterly J Econ 118:269–298
Caselli M, Falco P, Somekh B (2024) Inside the NBA Bubble: how Black players performed better without fans. J Population Econ 37:39
Chalfin A, McCrary J (2018) Are US cities underpoliced? Theory and evidence. Rev Econ Statistics 100:167–186
Choi H, Varian H (2012) Predicting the present with Google Trends. Econ Record 88:2–9
Conneau A, Khandelwal K, Goyal N, Chaudhary V, Wenzek G, Guzmán F, Grave E, Ott M, Zettlemoyer L, Stoyanov V (2018) Unsupervised cross-lingual representation learning at scale. arXiv:1911.02116
Dawson P, Downward P, Mills TC (2014) Olympic news and attitudes towards the Olympics: a compositional time-series analysis of how sentiment is affected by events. J Appl Stat 41:1307–1314
Devlin J, Chang M-W, Lee K, Toutanova K (2018) Bert: pre-training of deep bidirectional transformers for language understanding. arXiv:1810.04805
Enikolopov R, Makarin A, Petrova M (2020) Social media and protest participation: evidence from Russia. Econometrica 88:1479–1514
Entman RM (1993) Framing: toward clarification of a fractured paradigm. J Commun 43:51–58
Erbring L, Goldenberg EN, Miller AH (1980) Front-page news and real-world cues: a new look at agenda-setting by the media. American J Political Sci 16–49
Farina E, Pathania V (2020) Papal visits and abortions: evidence from Italy. J Population Econ 33:795–837
Forrest D, Simmons R, Buraimo B (2005) Outcome uncertainty and the couch potato audience. Scottish J Political Econ 52:641–661
Gillitzer C, Prasad N (2018) The effect of consumer sentiment on consumption: cross-sectional evidence from elections. American Econ J: Macroecon 10:234–269

Gorodnichenko Y, Pham T, Talavera O (2021) Social media, sentiment and public opinions: evidence from\# Brexit and\# USElection. European Econ Rev 103772
Hamborg F, Donnay K, Merlo P (2021) NewsMTSC: a dataset for (multi-) target-dependent sentiment classification in political news articles. (Association for Computational Linguistics (ACL))
Hatte S, Madinier E, Zhuravskaya E (2021) Reading Twitter in the newsroom: Web 2.0 and traditional-media reporting of conflicts. Technical Report
Hausman JA, Leonard GK (1997) Superstars in the National Basketball Association: economic value and policy. J Labor Econ 15:586–624
Hendel I, Lach S, Spiegel Y (2017) Consumers' activism: the cottage cheese boycott. RAND J Econ 48:972–1003
Ho VA, Nguyen DH-C, Nguyen DH, Pham LT-V, Nguyen D-V, Nguyen KV, Nguyen NL-T (2020) Emotion recognition for Vietnamese social media text. In: Computational Linguistics: 16th International Conference of the Pacific Association for Computational Linguistics, PACLING 2019, Hanoi, Vietnam, October 11–13, 2019, Revised Selected Papers 16 (Springer, 2020), pp 319–333
Jun S-P, Yoo HS, Choi S (2018) Ten years of research change using Google trends: from the perspective of big data utilizations and applications. Technological Forecasting Social Change 130:69–87
Kanazawa MT, Funk JP (2001) Racial discrimination in professional basketball: evidence from Nielsen ratings. Econ Inquiry 39:599–608
Lamla MJ, Lein SM (2014) The role of media for consumers? Inflation expectation formation. J Econ Behavior & Organization 106:62–77
Lerner JS, Han S, Keltner D (2007) Feelings and consumer decision making: extending the appraisal-tendency framework. J Consumer Psychol 17:181–187
Liu Y, Ott M, Goyal N, Du J, Joshi M, Chen D, Levy O, Lewis M, Zettlemoyer L, Stoyanov V (2019) Roberta: a robustly optimized Bert pretraining approach. arXiv:1907.11692
Loureiro D, Barbieri F, Neves L, Anke LE, Camacho-Collados J (2022) Timelms: diachronic language models from Twitter. arXiv:2202.03829
McCombs ME, Shaw DL (1972) The agenda-setting function of mass media. Public Opinion Quarterly 36:176–187
Nielsen (2021) Nielsen. https://www.nielsen.com/
Reed WR (2015) On the practice of lagging variables to avoid simultaneity. Oxford Bullet Econ Stat 77:897–905
Scherer AG, Palazzo G, Matten D (2014) The business firm as a political actor: a new theory of the firm for a globalized world. Business & Soc 53:143–156
Scheufele DA (1999) Framing as a theory of media effects. J Commun 49:103–122
Shen D, Urquhart A, Wang P (2019) Does twitter predict Bitcoin? Econ Lett 174:118–122
Smith C, Korschun D (2018) Companies can't avoid politics-and shouldn't try to. Harvard Business Rev
Soroka S, Daku M, Hiaeshutter-Rice D, Guggenheim L, Pasek J (2018) Negativity and positivity biases in economic news coverage: traditional versus social media. Commun Res 45:1078–1098
Soroka SN (2006) Good news and bad news: asymmetric responses to economic information. J Politics 68:372–385
Stock JH, Wright JH, Yogo M (2002) A survey of weak instruments and weak identification in generalized method of moments. J Business & Econ Statistics 20:518–529
Vosen S, Schmidt T (2011) Forecasting private consumption: survey-based indicators vs. Google trends. J Forecasting 30:565–578
Zhuravskaya E, Petrova M, Enikolopov R (2020) Political effects of the internet and social media. Annual Rev Econ 12:415–438