# Online Word of Mouth and the Performance of New Products

Lachlan Deer *
Department of Economics, University of Zurich
and
Pradeep K. Chintagunta
Chicago Booth School of Business
and
Gregory S. Crawford
Department of Economics, University of Zurich & CEPR

June 27, 2019

**JOB MARKET PAPER**

For the latest version click here

**Abstract**

We investigate the effects of online word of mouth on the demand for new products using Twitter data. Twitter can both generate buzz & awareness as well as provide information on product quality that can readily diffuse through the population. Leveraging comprehensive data from the US movie industry and Twitter, we estimate a structural model of consumer demand for attending theatrical releases in 2014-2015 that incorporates both information channels. The results show that both channels are important, but differ across types of movies. We find pre-release tweet volume is the most important channel for large franchise movies, generating buzz that influences box office earnings on the opening weekend. Demand for mid tier movies responds to increasing awareness driven by the volume of tweets posted after a movie is released. In contrast, the sentiment expressed in online WoM after a movie's release influences box office demand in subsequent weekends for smaller movies.

*Keywords:* Online Word of Mouth, Twitter, New Products, Demand Estimation, Movies

*The authors gratefully acknowledge useful comments and suggestions from Julian Langer, Yong Liu, Christian Oertel, Ralph Ossa, Peter Schmidt, Rahul Telang and Charles Weinberg, and from conference and seminar audiences in EMAC 2019, Marketing Science 2018, Swiss IO Day 2018, EARIE 2018, Mallen20, Bocconi and Zurich. We thank Cristian Dagnino, Yuhan He, Eugene Miravete and Akhil Rajan for their invaluable research assistance at the early stage of this project. All errors remain our own. Author emails: Deer (Corresponding Author): `lachlan.deer@gmail.com`, Chintagunta: `pradeep.chintagunta@chicagobooth.edu`, Crawford: `gregory.crawford@econ.uzh.ch`

# 1 Introduction

Social media platforms are important players in the modern marketing landscape, with American consumers now spending an average of 135 minutes per day on social media (Ward 2018). Twitter has emerged as the dominant platform where consumers seek and provide opinions about brands and products (Smith et al. (2012), Jansen et al. (2009)). The content posted on Twitter provides a wealth of information that can both alter consumers' awareness, as well as their beliefs about brands' perceived quality. This is particularly important for new products, where awareness can be low and quality perceptions are often evolving. Consumers also use Twitter to post expressions of anticipation, which we call 'buzz', about a product's impending release which is perceived to be important for generating demand over their initial launch (Houston et al. (2018)). The rise in importance of Twitter and the 'user generated content' posted on the site has fostered widespread growth in the perceived importance of word of mouth (WoM) marketing, which is now considered to be one of the most effective forms of marketing by marketing executives and the business press (WOMMA 2013). This belief is in part rationalized by the fact that consumers view information sourced from peers as more trustworthy than marketing coming directly from firms (Nielsen 2013). Despite Twitter's importance as a platform for consumers to discuss brands and products, there is limited academic research quantitatively linking content posted on the site to market outcomes.[1]

Understanding the magnitudes and mechanisms linking online WoM to product demand is an important input into the design of effective product release and marketing strategies for new products. There are two likely channels through which Twitter and other social media can influence the demand for products. First, the *volume* of WoM can raise awareness when consumers see tweets about a product with which they are unfamiliar and via reinforcing consumer's exposure to previous marketing messages. Increases in tweet volume can also increase buzz surrounding a product by aggregating observable expressions of anticipation about a product's impending release to a single place.[2] Second, the sentiment encapsulated

[1]This shortage of research also exists at the firm level. Although marketing executives believe in the importance of social media, only 15 percent of Chief Marketing Officers as of 2013 were able to quantitatively show its impact on business outcomes (American Marketing Association 2013).

[2]In the empirical model outlined in this paper we will be unable to fully separate the awareness & buzz channels present within the volume of tweets. As a result, through the paper we will link the buzz and

in tweets about new products facilitates information diffusion as aspects of product quality are revealed to potential consumers by analyzing the sentiment expressed in tweets sent by influencers and peers they follow. Examples of industries that are most likely to be affected by these channels include movies, music, videogames, consumer electronics, and fashion.

This paper quantifies the impact of online WoM on the demand for new products. Prior studies have strongly supported the notion that online WoM is an important source of information for consumers when they make purchase decisions (Godes & Mayzlin (2004), Chevalier & Mayzlin (2006), Liu (2006), Chintagunta et al. (2010)). However, existing research has typically been hampered by endogeneity concerns due to the positive correlation between online WoM and unobserved external sources of information and/or weak instruments, both of which manifest themselves as large WoM elasticities. We adopt a rich econometric specification to partial out this endogeneity using fixed effects based on ex-ante expected sales and the number of days since a movie has been released interacted with important product characteristics. We also differentiate ourselves from the existing literature across two substantive dimensions motivated by our study of an industry with many, heterogeneous new products: (i) by allowing demand elasticities for volume and sentiment to vary over a product's release period, an important feature for new products where demand is constantly adjusting as new information enters the market, and (ii) by providing novel evidence emphasizing how the volume and sentiment of Twitter posts differentially impact demand across products with different attributes, revealing heterogeneity in the mechanisms though which online WoM can influence demand.

Our analysis measures the impact of online WoM on new products in the context of the US movie industry. Movie studios, like all producers of entertainment products, closely track the evolution of discussion about their new releases on social media by purchasing data from multiple social media listening companies. Twitter has become the central conversation platform for movie discussion, and the effect of tweets on movie demand is hotly debated in the industry press and among firms providing predictive analytics for the industry (Suslak (2014), Wong et al. (2012), Baek et al. (2017)). This debate is taking place as the industry's rhetoric is shifting from a mantra centered around William Goldman's

awareness concepts together, i.e. "buzz & awareness." We are actively working on developing a natural language processing algorithm to isolate 'product buzz' as an extension to this paper.

famous phrase "nobody knows anything" to building up a data driven understanding of the industry. Our results contribute to this debate, showing that the impact of Twitter on demand is quantitatively important and highlighting that the mechanisms linking it to demand differs across types of movies. The results also provide insights into the impact of WoM on demand for new products generally.

The focus on the movie industry, and in particular wide release movies, is driven by several factors that create empirical leverage to measure the impact of Twitter on demand.[3] First, the industry is characterized by the frequent entry of new products, with between two and four new theatrical releases each week. Movie box office earnings typically decline sharply over time, with an average of eighty percent of revenue earned in the first three weeks. Second, awareness, buzz and uncertain product quality are important characteristics in the industry. These characteristics are believed to be key drivers of the success of a movie over its release window - slowing the decay in sales over time. Movie studios try to build up buzz in the pre-release phase to generate consumer interest, anticipating that it translates into increased awareness, greater willingness to see a movie and ultimately box office success. Because movies are experience goods, even when consumers are aware of which movies are in cinemas, they are uncertain about their consumption utility and frequently rely on the experience of peers to gauge a movie's quality.

There are two components to our empirical analysis. The first is the construction of a comprehensive data set on the US movie industry. We collect all movie-relevant tweets for wide release films distributed in 2014 and 2015 to measure online WoM.[4] Our search criteria returned 48 million tweets for these movies, an average of over 300,000 tweets per film. The volume of tweets, while clearly important for measuring buzz and/or awareness, cannot capture consumer *sentiment* about a movie. Since peers' perceptions of a movie's quality may be an important element of WoM, we augment the volume of tweets with a measure of each tweet's sentiment using state-of-the-art natural language processing techniques to calculate a measure perceived movie quality that evolves over time.[5] The Twitter data

[3]The focus on wide release movies allows us to abstract away from modelling the strategic transitions from limited-to-wide release by movie studios.

[4]Song et al. (2019) show that consumer discussion on microblogs is a stronger predictor of demand than discussion on third party platforms.

[5]We use the VADER sentiment lexicon to classify tweets as positive, negative or neutral in sentiment (Gilbert & Hutto (2014)). The VADER lexicon has been documented to be the most consistent in terms of

are combined with additional data from numerous sources to provide a comprehensive overview of the industry. These additional data include daily national box office revenue, characteristics of a movie at the time of its release and production budgets from Box Office Mojo, pre-release expectations of box office earnings from the Hollywood Stock Exchange, critic reviews from Metacritic, opening night consumer reviews from CinemaScore and detailed advertising data from Nielsen AdIntel.

This data is used to estimate a structural model of movie industry demand. The key outputs are estimated demand elasticities for tweet volume and sentiment which we allow to vary over time and across types of movies (e.g. sequels vs non-sequels). Formally, we build a nested logit demand model where consumers decide each week whether to see a movie in theatres and if so, which one. The model allows the volume and sentiment elasticities to differ between a movie's opening weekend and post-opening weekends, capturing the differential importance of buzz, awareness and sentiment across a movie's release window. Increasing levels of buzz and awareness, measured in our model by tweet volume, are expected to translate into consumers' viewing the movie immediately upon release. On the other hand, tweet sentiment is expected to have a greater influence on demand after the opening weekend, i.e. after early viewers have seen the movie and post reviews. The demand model also accommodates advertising by studios, competition between films at the box office, declining consumer valuations for a movie over the release window, and the frequent rotation of movies through cinemas.

A key ingredient in the demand model is a comprehensive set of fixed effects that capture consumer awareness, interest and quality perceptions about different types of movies that stem from offline WoM and any other sources of information. In particular, we include fixed effects based on ex-ante expected box office performance, opening night consumer reviews and time varying fixed effects for movie sequels and for each movie genre. Ex-ante expected box office and sequel fixed effects that vary across the number of days a movie has been released are included to soak up increased consumer awareness and interest for movies with higher expected earnings and/or for movie franchises. Opening night consumer reviews are

performance, when compared to alternative lexicons *and* standard machine learning algorithms, for social network and online comment data when using a three category categorization of positive, negative and neutral (Ribeiro et al. (2016)).

integrated to partial out quality perceptions that may be shared offline. Time-varying fixed effects for each movie genre capture consumer awareness and quality perceptions that persist over genre and may evolve differently over a movie's release. Absorbing variation from external sources through this rich set of fixed effects allows us to isolate the impact of online WoM on demand.

Across all movies, our results reveal an average elasticity of market share with respect to pre-release volume of 0.05 on the opening weekend, an elasticity of 0.08 for post-release volume after the opening weekend, and a post-release sentiment elasticity of 0.27. These estimates are substantially lower than the existing estimates of the impact of online WoM on new product demand, which find large volume elasticities ranging from 0.57 to over 1 (Liu (2006), Kim & Hanssens (2017). Dhar & Chang (2009)) and sentiment elasticities around 0.4 (Gopinath et al. (2013)).[6] The large difference in magnitudes provide suggestive evidence that our fixed effect strategy was successful in cleaning out key endogeneity concerns. Furthermore, our estimated volume elasticities align closely to recent work studying the elasticities viewership of *existing* TV shows to the volume of online WoM which find estimated elasticities between 0.02 and 0.04 (Seiler et al. (2017), Lovett & Staelin (2016)). Unlike these findings for TV viewership, we find that sentiment is an important determinant of demand for movies.

These across-movie averages, however, mask important differences in the relative importance of different information channels on demand across three different types of movies. First, for franchise movies we find that the pre-release volume of WoM facilitates consumer buzz & anticipation that in turn drives consumer demand, with the effect being strongest on a movie's opening weekend.[7] Across all such movies we find an opening weekend demand elasticity of 0.17 with respect to pre-release tweet volume. This effect is even stronger for the large franchise movies which have high pre-release expectations of success, where the estimated elasticity rises to 0.34. Second, for smaller movies our results indicate that

[6]Recent meta-analysis by You et al. (2015) reveal average volume elasticities close to 0.2 and a sentiment elasticity of 0.4 across a range of both new and existing products. Existing work usually relies on numerical rating scales provided within a blog entry to form a measure of the *sentiment* of a review. Numerical ratings are typically termed 'valence' in the existing literature whereas sentiment is used to describe a measure of positivity in text. Throughout the paper we will use the term 'sentiment' to refer to both numerical and text based measures of quality to ease exposition, and mention whether the measure is based on numerical ratings or text where important.

[7]Throughout the paper we will use the terms 'sequel' and 'franchise' interchangeably.

online WoM plays an key role in facilitating information transmission and social learning about a movie's perceived quality. Our estimates reveal a post-opening demand elasticity with respect to sentiment of 0.7, again more than double the effect on the average movie. Positive changes in sentiment after a movie's release lead to positive updates in quality expectations that translate into economically large and significant increases in demand in subsequent weeks. Finally, for ex-ante mid-tier movies our results highlight that the volume of tweets in the post-release phase are an important determinant of demand, with an elasticity estimate of 0.23. This suggests a mechanism of expanding consumer awareness about these movie's release in the weeks *after* opening is important for generating box office success. Taken together, these results are the first in the literature to reveal how the mechanisms of online WoM have heterogeneous impacts on the demand for products with different product characteristics.

The remainder of this paper proceeds as follows. Section 2 reviews the relevant literature. Section 3 presents an overview of the movie industry and the importance of Twitter based social media in it. In Section 4 we describe a conceptual framework linking Twitter volume and sentiment to consumer demand and outline how our Twitter data maps into these measurements. Section 5 outlines our model of movie industry demand and details the estimation strategy. Parameter estimates & demand elasticities are presented in Section 6. Section 7 discusses the marketing implications of our findings and concludes the paper.

## 2 Related Literature

We contribute to the substantial existing literature that seeks to evaluate the effect of online WoM on product demand. A comprehensive meta-analysis by Babić Rosario et al. (2016) finds that electronic WoM has the largest impact on new products, and documents that volume has a stronger impact on sales than sentiment. A central feature of the existing literature that studies the impact of online WoM on new product demand is that identifies the effects of WoM through timing assumptions, usually assuming that WoM at time $t-1$ effects new product sales at time $t$ (Godes & Mayzlin (2004), Liu (2006)). This has also been the predominant approach to estimating the effects of WoM in the movie industry, using movie blogs to measure WoM. Chintagunta et al. (2010) and Gopinath et al. (2013) combine this strategy with the sequential release of movies across DMAs to understand the impact of advertising and movie blog posts on demand. Their findings emphasize the role of pre-release advertising and blog volume on opening night sales, and blog sentiment - as measured by numerical ratings - on post-opening box office. Kim & Hanssens (2017) studies how blog volume and advertising impacts pre-launch consumer interest surrounding movies and video games, finding that blog posts generate permanent, trend-setting effects while advertising only causes temporarily increased interest. We contribute to this literature in three ways. First, we provide an in-depth study of WoM using Twitter data, as opposed to blog posts, an important extension given the emergence of Twitter as the dominant platform for new product discussion. Second, we advance the measurement of the sentiment of online WoM by using state of the art sentiment dictionaries to infer it from written text, rather than relying on numerical ratings. Finally, our empirical approach is designed to prevent our measures of online WoM capturing the effects of offline WoM resulting in WoM elasticities that are significantly lower than existing estimates. We are also the first to emphasize substantial heterogeneity in the effect sizes across movies with different characteristics.

Our work also relates to an emerging literature seeking to understand the role of social media on consumer demand for entertainment products. To date, this literature has focused on the movie and TV industries. Hennig-Thurau et al. (2015) and Gelper et al. (2018) investigate the impact Twitter discussion on movie demand, focusing on opening weekend box office outcomes. Hennig-Thurau et al. (2015) highlights an association between negative tweets posted on opening night and a faster drop in earnings over the opening weekend. Gelper et al. (2018) study the role of spikes in the volume of conversation on opening weekend box office revenues revealing that the number of jumps in conversation in blogs and on Twitter are positively correlated to opening sales. Numerous studies provide correlational evidence linking volume and sentiment of tweets to box office revenue using a range of sentiment classifiers (Duan et al. (2008), Apala et al. (2013), Jain (2013), Rui et al. (2013), Baek et al. (2014), Ding et al. (2017)). We extend these results in two important ways. First, we extend results in Hennig-Thurau et al. (2015) and Gelper et al. (2018) by estimating the impact of tweet volume and sentiment over three weekends of release - allowing for buzz, awareness, and information diffusion to impact movie revenues beyond the opening weekend. Second, we build a comprehensive data set containing all pre- and post- release tweets for two years of movies instead of the small sample of tweets made available from Twitter's public API or from screen scraping. Finally, our results demonstrate differential impacts of social media across different movie types.

Existing work studying relationship between TV show viewership and social media generally utilizes data from the Chinese microblogging platform Sina Weibo. Gong et al. (2018) finds that retweets of informative content by influencers and studio's own tweeting impacts subsequent TV series viewership. Seiler et al. (2017) investigates the impact of online WoM using a natural experiment generated by a temporary block in access to Weibo imposed by the Government on mainland Chinese citizens. Their findings highlight the importance of the volume of posts and document that sentiment has no causal effect on demand, emphasizing a mechanism of social consumption. Our approach relates to these studies in two ways. First, our focus on movies provides a platform to study new products rather than existing ones. This is important given WoM's increased relevance for new products. Second, the mechanisms through which WoM influences new product demand are different than those for existing products. Our results emphasize the role of buzz, awareness and social learning as opposed to social consumption.[8]

# 3 The US Movie Industry & Twitter

## 3.1 The US Movie Industry

United States box office revenue averages over 10.5 billion dollars per year since 2010, with over 1.2 billion ticket sales. The US market is the largest in the world, with total US Box Office revenue accounting for approximately 25 percent of global sales (MPAA 2017). Each year approximately three quarters of the US and Canadian population go to the cinema

[8]There is a recent debate in the economics literature about the role of social consumption and movie demand. Gilchrist & Sands (2016) use local weather shocks to instrument for WoM and approximate DMA level box office using Google Trends data to study network externalities in movie consumption. finding a role for both social consumption and social learning. Kuehn & Lampe (2018) documents that Gilchrist & Sands (2016) are sensitive to data quality and empirical specifications, downplaying the role of social consumption. Compared to these papers, we use observable measures of WoM - both offline and on Twitter - to a estimate the response of demand.

at least once, and 35 percent see at least six movies per year (MPAA 2016). Furthermore, success in the US theatrical window is important for the overall profitability of a film. A strong US box office often serves as a signal of quality in foreign markets and is strongly correlated with DVD and streaming sales.

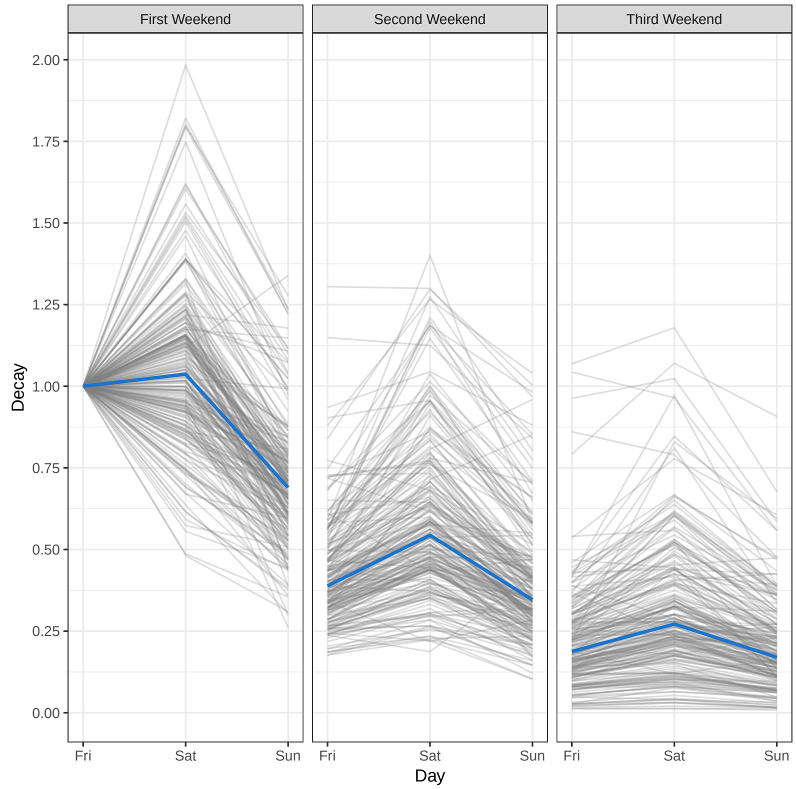
Movie releases are categorized based by the size of their initial release. Wide release films are typically released across the US on the same day, usually Friday, and are intended to have wide audience appeal. These films include blockbuster titles like *The Martian* and include sequels and franchise films like *Star Wars* and *The Hunger Games*. Limited release films open to a smaller number of cities and often have a more niche audience. Depending on the its initial success and critic response, a limited release film may then roll out gradually across the entire country. We omit limited and limited-to-wide release films for our analysis to abstract away from the film studios' release strategy decision.[9]

Individual movies' theatrical life cycles are typically short lived. Each week 3-4 new movies enter and a similar quantity exit the market. This rotating menu of movies means consumers may not be aware of which movies are currently showing in cinemas. Each movie is unique and has an ex-ante uncertain quality. This quality uncertainty means that before going to see a movie, consumers are unsure as to whether they will like it. To overcome this uncertainty they make their movie viewing decisions based on observable characteristics of the movie - such as actors, genre, production budget and MPAA rating, as well as other information sources such as advertising, critic reviews and word of mouth from other consumers. When consumers look to Twitter to gauge WoM from peers, both the volume of tweets and the sentiment expressed are potentially relevant sources of information.

The focus of this study is the box office performance of wide release movies over their first 3 weekends of release. Nationwide Box Office earnings for wide release movies released in 2014 and 2015 are obtained from Box Office Mojo. Over our study period, 86 percent of a movie's total box office gross is earned inside its first three release weeks. Furthermore, 80 percent of the realized box office revenue over these three weeks is taken on weekends, Friday to Sunday. There is a large degree of heterogeneity in a movie's box office earnings

[9] We also abstract away from the release timing decision of studios. For wide release films we think this is not of first order importance because release dates are typically set well in advance of the actual release date.

Figure 1: Box Office Decay over the Opening 3 weekends




*Note*: This figure depicts the decay in box office revenue over the opening three weekends of release compared to opening night receipts ($t = 0$). Decay of movie $j$ on day $t$ is then defined as $\text{decay}_{jt} = \text{box office}_{jt}/\text{box office}_{j0}$. The grey lines plot decay patterns of each wide release movie released over 2014-15. The bold blue line plots the median decay over time, where the averages are computed at the daily level.

both on its opening night and over the duration of its theatrical release. The median film's opening night box office is $14.6 million USD, with an interquartile range between from $8.1 to $30.3 million dollars. This heterogeneity extends to total US box office earnings, with median earnings of $44.5 million and an interquartile range going from $19.1 million to $90.2 million. Box office success persists over a movies release, the correlation between opening night box office and gross box office is over 0.9.

Figure 1 documents how a film's box office earnings decay over time relative to its' opening night.[10] The blue line in Figure 1 plots the median decay over time, revealing that

[10]A movie's opening night is generally the day with the highest box office earnings, and thus a good benchmark to evaluate the 'stickiness' in the time path of box office earnings. We define Box Office decay of a movie $j$ as the ratio of box office revenue on day $t$ compared to box office revenue on the opening night

box office revenue declines quickly over time, with the median movie's box office on the third Sunday of release falling to 17 percent of opening night earnings. There is a large divergence in sales trends across movies, with some movies still earning daily revenues above half their opening night by the third Sunday, while others are closer to zero. The level of decay is exhibits strong autocorrelation, with movies that decay slower in earlier weekends maintaining this pattern over time. Within a weekend, there is a hump shape in decay, indicating box office is on average higher on Saturday, than either Friday or Saturday. This is because there are more screenings that individual consumers prefer to attend on a Saturday than on the other days.[11]

*Additional Movie Industry Data.* In addition to the box office sales data described above, we collect detailed additional data about each movie to construct a comprehensive view of the movie industry. A brief summary of the data is located in Table 1, showing the data sources and measures constructed from each source. A more detailed description of the data sources and variable definitions is contained in Appendix A.

## 3.2 Twitter

Twitter is a microblogging website that launched in July 2006. It is the largest microblogging platform on the internet and is in the top 15 most visited websites globally since 2013. By 2012/13 Twitter had more than 40 million monthly active users in the United States and 100 million worldwide and continued to grow to 300 million active users by 2016.

The main action users of the Twitter platform take is to post short excerpts of text, i.e to 'tweet', voicing an opinion or sharing information.[12] A second important action is "retweeting." Retweets allow users to forward the tweets of other users to their own followers - expanding the visibility of the original tweet to a new audience. When retweeting, users can optionally comment on the tweet that they are forwarding. In addition to tweeting

$(t = 0)$, $\text{decay}_{jt} = \text{box office}_{jt}/\text{box office}_{j0}$. This measure of decay does not force movie revenues on day $t$ to be less than opening night box office, so $\text{decay}_{jt}$ is not constrained to lie below 1. As Figure 1 shows, often on the opening Saturday a movie earns more than on the opening night.

[11] Popular Friday show-times are in the evening, after work. On Saturday popular show-times extend from the early afternoon late into the evening. Then on Sunday, popular show-times are Sunday afternoon, with many consumers "staying in" on Sunday evenings before work begins on Monday. Saturday afternoon screenings may also be more attractive to families.

[12] Over the 2013-2016 time period tweets were restricted in length to a maximum of 140 characters of text, and one multimedia element.

Table 1: Additional Movie Industry Data

| Data Type | Data Source | Measurement |
|---|---|---|
| Advertising Spend | Nielsen Ad Intel | Pre-release Ad Spend<br>Post-release Ad Spend |
| Movie Characteristics | Box Office Mojo | Release Date<br>Genre<br>Franchise Indicator<br>Lead Actors |
| Production Budget | Box Office Mojo<br>The Numbers<br>Wikipedia | Production Budget ($ millions) |
| Actor Starpower | Oscars Database | Nominations and Wins<br>of Academy Awards |
| Critic Review<br>Offline Consumer Review | MetaCritic<br>CinemaScore | Metascore<br>CinemaScore grades |
| Expected Box Office | Hollywood Stock Exchange | HSX Closing Price |

*Notes:* This table shows the sources of data used to construct a movie characteristics. We define a movie to be part of a franchise if the movie is a sequel of another - where plot depends on the previous movie, or is part of a larger movie franchise where story lines are not necessarily dependent on previous movies. We include 9 genres: drama, sci-fi/fantasy, action/adventure, romance, comedy, family, thriller, horror and 'other.' Movies are allocated to genres using first mentioned genre on a movie's Box Office Mojo webpage. Actor starpower for an actor, $a$, starring in a movie in year $t$ is defined as $\text{Actor Starpower}_{at} = \sum_{\tau<t} \text{Nominations}_a + 2 \times \sum_{\tau<t} \text{Awards}_a$. The starpower measure for a movie, $j$ is then the sum of these scores across lead actors, $\text{starpower}_{jt} = \sum_{a \in \mathcal{A}} \text{Actor Starpower}_{at}$. CinemaScore grades are from opening night reviews of consumers collected by CinemaScore at theatres. There are 10 CinemaScore grades ranging from A+ to D. The Hollywood Stock Exchange is an online prediction market that predicts a movie's box office revenue over opening four weeks. Spann & Skiera (2003) and **?** demonstrate that HSX trading prices close to the release date provide reliable forecasts of predicted box office performance. Stocks trade at market prices that reflect the market's collective expectation. For example, if a particular movie stock trades at H$60.00, the market is predicting that the movie will gross US$60 million at the box office in the first four weekends of wide release. Further details about the variable definitions are contained in Appendix A.

and retweeting, users "follow" each other, subscribing to their tweets. The tweets of users that an individual chooses to follow then automatically appear on the individual's feed in chronological order.[13] While Twitter was envisaged as a platform for users to tweet and retweet, 44 percent of subscribed users have never posted a tweet, instead using it as a platform to consume information (Murphy 2014). Twitter handles more than 1.6 billion search queries per day - where a user searches a topic to find relevant tweets, similar to a Google search (Siegler 2011).

Movies are one of most discussed topics on Twitter with over 200,000 movie relevant tweets posted per day (Hu et al. (2017), Suslak (2014)). This is because Twitter's core demographic, 18-29 year-old urban residents, is a core audience for movies. A recent Nielsen study revealed that Twitter users are 340 percent more likely than non-users to have seen more than 12 movies over a twelve month period and are twice as likely to have seen a movie over the first 10 days of release (Jarvey (2014), Robehmed (2015)). Among the Twitter users surveyed there was strong evidence that they view the platform as an important platform for movie discussion. Sixty three percent revealed that they heard about upcoming and in-release movies through social media. Eighty-seven percent of surveyed users stated that their recent movie attendances were influenced by the tweets of others. Over half of movie-going Twitter users post their thoughts about a movie after they leave the theatre.

# 4 Conceptual Framework & Measurement

This section outlines a conceptual framework detailing how the volume of tweets and sentiment expressed within them can influence consumer demand. We explicitly link these observable variables to the underlying economic mechanisms at work. After discussing the mechanisms, we outline the collection of the Twitter data linking these concepts to the data at hand.

[13]In early 2016, Twitter shifted from a chronological ordering to an algorithmic ordering that sorts tweets by quality and timeliness. As of late 2018, Twitter brought back the option to view tweets in chronological order.

## 4.1 Framework

**Tweet Volume.** The volume of tweets can influence the demand for movies through two mechanisms: increasing product awareness and creating 'buzz' about a movie.[14] Twitter posts can raise awareness about a movie if consumers who previously did not know about it see (or search) for posts that contain information about it. This could be through seeing a movie trailer posted by the studio or retweeted by another consumer or other information posted in a tweet by someone they follow. In addition to increasing awareness directly, tweets can increase awareness by reinforcing previous advertising a consumer has seen. Movie studios are large advertisers, but there is evidence that consumers forget about the advertising messages they see and that reinforcement through multiple exposures is important for enhancing awareness (Zielske (1959), Campbell & Keller (2003)). Seeing tweets about a particular movie while scanning through their Twitter feed can then remind or reinforce a message to a consumer that a movie is about to be released or is already showing in cinemas, which can lead to an increase in probability that a consumer will go to the cinema to see the movie.

The volume of tweets can also impact the demand for a movie due to product 'buzz.' We define product buzz as the aggregation of observable expressions of anticipation about a new product. These expressions are usually observed prior to a product's release or over the period of its initial launch. Buzz is thought to be an important determinant in demand for experience goods that have a strong social consumption component (Houston et al. 2018). Buzz about a movie increases demand because consumers enjoy seeing movies that their peers have seen and are talking about. This effect of buzz manifests in the volume of tweets because these anticipatory comments are posted on Twitter by consumers who want to act as opinion leaders, and reflect their interests, excitement, and expectations (Toubia & Stephen (2013), Hennig-Thurau et al. (2004), Sun et al. (2006), Chu & Kim (2011), **?**). These type of posts are typically neutral in sentiment. Whilst Houston et al. (2018) emphasize the pre-release nature of buzz and its' importance for initial sales, our framework allows buzz to influence sales over a longer horizon because we include measures

[14] With our current data we are unable to disentangle the effects of awareness and buzz. We are actively pursuing a text analysis algorithm which aims to isolate consumer buzz, which we can then use to separate these effects.

of the volume of pre- and post-release tweets in our demand model for the first three weekends of a movie's release. Buzz may impact demand over a longer horizon because anticipatory emotions expressed in tweets about seeing a movie can spill over to influence demand in subsequent weeks, particularly as not all consumers can attend a movie on its opening weekend.

We acknowledge that Twitter is not the sole source of product buzz and consumer awareness for new movies. Consumers can interact on other online platforms or offline by talking to their friends and work colleagues about upcoming movies. Awareness and buzz can also be driven by the advertising of movie studios. In our empirical exercise we control for these aspects in order to capture the impact of the volume of online WoM, as measured by tweets, rather than *general* WoM for that tweet volume serves as a proxy variable. Further details outlining our strategy for accounting for offline WoM are postponed to the discussion of our empirical specification in Section 5.1.

**Tweet Sentiment.** Before and during a movie's release the volume of tweets jointly measures consumer awareness and movie buzz, but it cannot measure consumer perceptions of a movie's quality. The sentiment expressed in the text of a tweet *can*, however, provide a measure of these perceptions. Sentiment can impact movie sales because it allows information to diffuse through the population, providing an avenue for social learning. Social learning occurs when aspects of movie quality are revealed to potential consumers through interactions with their peers.[15] Twitter provides a channel for social learning to occur because consumers who see a movie and are active on Twitter typically post short reviews or opinions after viewing (Jarvey (2014), Robehmed (2015)). Potential consumers can use these posts to gauge a movie's quality by analyzing the sentiment expressed in them before deciding whether to go to the movies and, if so, which movie to attend (Goldsmith & Horowitz (2006)). Over a movie's release these reviews and opinions grow in volume and a consensus opinion about movie quality emerges which can differ from pre-release expectations. This change in quality perceptions can then impact demand in subsequent weeks.

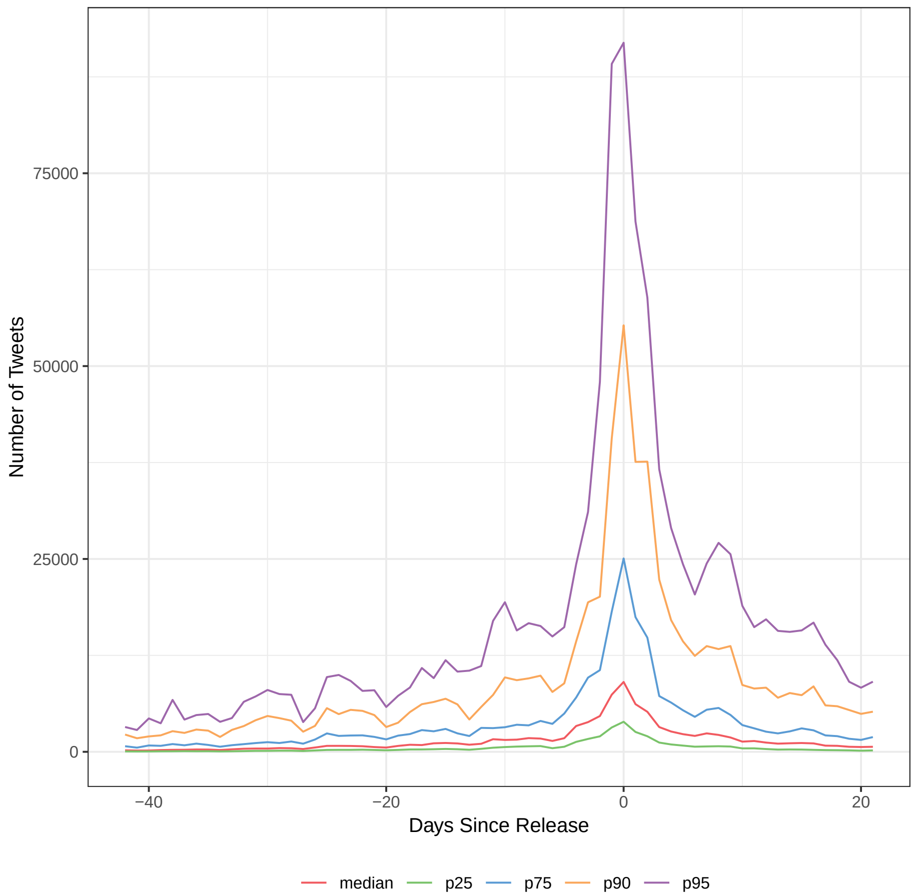
[15]Throughout the paper our discussion of movie quality refers to *vertical* quality which is common to all consumers, rather than horizontal match quality.

As with tweet volume, consumers can also obtain information about movie quality from sources other than Twitter, either from consumer review aggregators or offline discussion. Expert opinion is also available through critic reviews which are typically correlated with consumer's opinions . Because the goal of our analysis is to quantify the impact of sentiment in online WoM, we need to control for this potential confounding effect. In our empirical analysis we include critic reviews and a movie's CinemaScore - a measure of consumer review collected on a movie's opening night - to capture other offline and non-Twitter based WoM to control for these external sources of information.

## 4.2 Measurement

We retrieve movie-relevant tweets from Twitter's GNIP Historical Powertrack, which archives the complete history of tweets since Twitter's launch in 2006. Movie specific searches were constructed for all wide release films released in 2014 and 2015 to query the Twitter database and extract tweets related to each movie. In total 310 unique search queries were assembled. Searches were constructed to include the movie name, widely known abbreviations, movie-relevant hashtags and, where relevant, terms related to movie sequels.[16] The searches were restricted to return only English language tweets.[17] We use tweets posted from 60 days pre-release until the end of the third release weekend.[18] This is the period with the greatest volume of movie tweets that are relevant to a movie's theatrical release, and when our proposed mechanisms are likely to be influencing consumer demand. The searches returned approximately 50 million tweets over the 77-day search period, approximately 300,000 per movie.

Figure 2: Volume of Tweets around the Release Date




*Note*: Figure plots the temporal pattern of tweets across different percentiles of the distribution. The volume of tweets on a given day includes all original tweets and retweets about a movie on that day.

#### 4.2.1 Measuring the Volume of Tweets

Figure 2 documents the temporal pattern of tweets over the 77-day window running from 60 days pre-release until the end of the third week of release. On each day the number of tweets is the sum of original tweets posted and the number of times an existing tweet was retweeted by another user on a given day. Two clear observations emerge. First, there is large amount of heterogeneity in tweet volume across movies in the upper half of the distribution. Movies in the 90th percentile and above consistently have more than 5 times the tweets per day then the median movie in our data. Movies in the upper percentiles are typically blockbuster franchise movies, which for 2014-2015 include *The Hunger Games: Mockingjay*, *Jurassic World*, *Spectre*, and *The Avengers*. The second pattern is the the evolution in the number of tweets per day. Tweet volume increases pre-release in two steps. The first increase occurs approximately two weeks prior to release with a jump in the daily level of posts. This increase is most pronounced for films in the upper tail of the volume distribution. Tweet volume then rises markedly in the week immediately prior to release - peaking on the release day. Post-opening night, tweet volume decreases rapidly, with small 'bumps' on the weekends when more consumers are going to the cinema.

In the demand model that follows in Section 5 we split the volume of tweets into pre- and post-release measures. This allows us to assess whether the combined importance of buzz and awareness differ across stages of a movie's release. We measure pre-release volume of tweets for a given movie $j$ as:

$$prevol_j = \sum_{\tau=-60}^{-1} \text{tweets}_{j\tau}$$

and for post-release volume on day $t$ we sum all tweets from the date of release until day

[16] A complete list of the searches conducted are available in our Web appendices.

[17] Some initial searches proceeded without the English language restriction. We found that over 80 percent of tweets that were returned with our search rules were in English. Non-English language tweets were returned when a tweet was written in a different language but used an English language hashtag or '@ mention.' We chose not to geo-localize searches to only those known to be posted from the USA because (i) Twitter users can easily view Tweets posted from other countries, and (ii) only a small fraction (typically less than 5 percent) of Tweets provide geolocation data.

[18] The initial time window for the tweet retrieval was, for each movie, from 6 months pre-release until 6 months post-release

$t-1$:

$$postvol_{jt} = \sum_{\tau=0}^{t-1} \text{tweets}_{j\tau}$$

Throughout our analysis we define a day $t$ as the 24 hour period from 12.00pm to 11.59am, i.e noon until noon, on the US East Coast.[19] The variable $\text{tweets}_{j\tau}$ includes all original tweets and retweets about a movie on a given day. Pre-release volume is then the sum of all tweets from 60 days pre-release until midday on the Friday of release. The post-release tweet volume of a given movie on Saturday of the second weekend of release is the sum of all tweets and retweets made from midday on the opening Friday until midday of the second Saturday.

The inclusion of retweets is important for our conceptual framework for two reasons. First, consumers may become aware of a movie through seeing a retweet rather than the original tweet itself. Second, many Twitter users share content that reflects their views and feelings of anticipation via retweets instead of posting a new tweet themselves. The inclusion of retweets also provides us a simple means to account for influential Twitter users - their tweet is counted multiple times.

#### 4.2.2 From Tweet Content to Tweet Sentiment

We use the VADER sentiment lexicon to classify tweets as positive, negative or neutral (Gilbert & Hutto (2014)). The VADER lexicon is well suited for our task because it was specifically designed to analyze the sentiment of text on microblogs and social media. VADER incorporates word capitalization, punctuation, internet slang and emoticons into the sentiment scoring process - all of which are common components of Twitter language.[20] VADER's performance compares favorably to standard machine learning approaches across

[19]The decision to shift our definition of a day comes from two factors: (1) Most movie show times are scheduled from the afternoon onwards, (2) Twitter posts from movie goers who attend late night screenings often are posted after midnight on the East Coast, i.e after 9pm on the West Coast, and are relevant for potential consumers the next day. Further shifts to account for differing time zones across the US are not possible because box office earnings are reported daily at the national level. Our results are qualitatively similar if we define a day from 12.00am to 11.59pm on the US East Coast.

[20]These features are not incorporated into other established lexicons like LIWC, ANEW and Senti-WordNet. VADER's sentiment classification has been validated by humans and there is a high degree of correlation between the lexicon's classification and that of human encoders, with a correlation of 0.88 (Gilbert & Hutto 2014). Appendix B provides additional detail on the inner workings of the VADER sentiment lexicon.

a wide variety of contents including social media, movie reviews, product reviews and newspaper editorials.[21]

The VADER lexicon assigns a sentiment score for each tweet in the range $[-1, 1]$ by analyzing the words and syntactic structure contained in the text.[22] We use the sentiment score to classify each tweet into one of three categories: positive, negative or neutral using the following rule:[23]

$$\text{sentiment} = \begin{cases} \text{negative} & \text{if VADER score} < -0.05 \\ \text{neutral} & \text{if VADER score} \in [-0.05, 0.05] \\ \text{positive} & \text{if VADER score} > 0.05 \end{cases}$$

For each movie-day pair we compute the total number of tweets in each of the three categories.[24]

The VADER classified movie-day counts are then aggregated to compute the positive-negative ratio for movie $j$ on day $t$:[25]

$$sentiment_{jt} = \frac{\sum_{t=\tau}^{t-1} \#\text{ positive tweets}}{\sum_{t=\tau}^{t-1} \#\text{ negative tweets}}$$

This ratio-based measure of sentiment is the most common way to use classified text to

[21] Ribeiro et al. (2016) conduct a systematic evaluation of the relative performance of eighteen sentiment lexicons and machine learning algorithms on multiple data sets. They show the VADER lexicon is the most consistent in terms of performance for social network and online comment data when using a three category categorization of positive, negative and neutral.

[22] As for tweet volume we run the sentiment scoring on the combination of original tweets and retweets. See Appendix B for a detailed discussion of the scoring of individual tweets. Tweets with a score of less than 0 contain more negative words than positive. Tweets with a score of greater than 0 contain relatively more positive words. The closer a tweet is scored to one of the boundaries, the more polarized is the content - i.e the stronger is the positive or negative sentiment.

[23] The empirical results that follow are robust to changes in this threshold value. We have explored up to increasing the threshold such that the size of each bin being equal - i.e thresholds of $\pm 1/3$. The choice of 0.05 follows from recommendations prescribed by the Lexicon's developers.

[24] Computation of the tweet level VADER scores and the number of tweets in each category per movie day was performed on a cloud cluster using the University of Zurich's Science Cloud architecture. The cluster was initialized using ElastiCluster (Baer et al. (2018)) with 4 workers each having 32 GB of RAM distributed over 8 cores. Twitter's JSON data was stored in a OpenStack Object Store (OpenStack 2018) and loaded into Apache Spark (Zaharia et al. 2010, Zaharia et al. (2016)) using Alluxio, a Virtual Distributed File System (Li et al. 2014, Li (2018)). Computation was implemented using PySpark and managed using the Snakemake workflow management system (Köster & Rahmann 2012).

[25] We compute these ratios using the noon-until-noon definition of a day, consistent with how we measure tweet volume.

create a continuous measure of sentiment based on the opinions of social media users.[26]

We compute separate measures of pre- and post-release consumer sentiment. Splitting the sentiment measures into pre- and post- release allows us to account for potentially different strengths in the signal that sentiment encapsulates in these phases.[27] The separation also allows us to assess social learning. Conditioning on pre-release sentiment means that an increase in post-release sentiment reflects a positive update in movie quality. We can then trace out this effect on demand. Pre-release sentiment is computed using all tweets from 60 days pre-release up to the day immediately preceding release. The post-release measure of sentiment varies daily, so that post-release sentiment on day $t$ is computed using all post-release tweets up to and including $t-1$.

# 5 A Sliding Window Nested Logit Model of Movie Demand

We now outline a nested logit model of consumer demand for movies. The model allows us estimate elasticities that capture demand responses to the volume and sentiment of posts made on Twitter by making explicit assumptions about consumer preferences and competition that match stylized facts from the movie industry.[28] The model analyzes individual consumers' making decisions about their movie attendance using a Sliding Window Nested Logit (SWNL). The SWNL model augments the conventional nested logit model from Berry (1994) to account for the frequent rotation of movies in and out of cinemas (Ainslie et al. (2005)). The model of consumer decision making is then aggregated to the level of national market shares to match the information present in our data.

There are $i = 1, 2, 3, \ldots, M$ individuals on each day $t = 1, 2, 3, \ldots T$ deciding between attending one of the $j$ movies currently showing in cinemas or staying at home. Over $T$ days, $j = 1, 2, 3, \ldots, \mathcal{J}$ movies rotate through the cinemas. We denote the decision to stay at home and not see a movie as $j = 0$. Each movie $j \geq 1$ is released in period $r_j$ and is

[26] See, for example, O'Connor et al. (2010), Bollen et al. (2011), Nguyen et al. (2012) and Si et al. (2013).

[27] In the pre-release phase there is little movie quality relevant information that potential movie viewers could extract from tweets because no-one has seen the movie.

[28] See Appendix C for reduced form regressions motivating the structure of the demand model.

shown in cinemas for a consecutive number of days. In any period $t$, we denote $J_t$ as the set of movies that are currently showing. We assume that on day $t$, consumers know all movies, $j \in J_t$ that are available in cinemas and have access to a cinema showing each movie.

Each movie $j \geq 1$ is characterized by a set of movie characteristics, $\mathcal{X}_j = [w_j, x_{jt}, d_{js}, d_{jg}, \xi_{jt}]$.[29] $w_j$ is a $1 \times L$ vector of time-invariant movie characteristics, such as production budget, whether a movie is part of a franchise, critic reviews, and CinemaScore grade. $x_{jt}$ is a $1 \times K$ vector of movie characteristics that vary over the release window, which in our specification will be pre- and post-release tweet volume, tweet sentiment and advertising spending. Movies are categorized to belong to an expected performance tier $S$ (EPT), representing expected profitability, and genres $G$, representing their main thematic subject.[30] $d_s$ is a dummy variable taking the value 1 if movie $j$ belongs to EPT $s \in 1, 2, 3, \ldots, \mathcal{S}$ and $d_g$ is a genre dummy variable taking the value 1 if movie $j$ belongs to genre $g \in 0, 1, 2, 3, \ldots, \mathcal{G}$. EPTs $s$ and genres $g$ are each mutually exclusive and exhaustive, so movie $j$ belongs to one tier and one genre. $\xi_{jt}$ is a scalar, time-varying characteristic that is unique to each movie.[31] All movie characteristics are known by, and provide utility to, each consumer. As econometricians, we observes all movie characteristics except $\xi_{jt}$.

Consumer $i$'s conditional indirect utility of attending movie $j \in J_t$ is modeled as a linear function of product characteristics:

$$u_{ijt} = x_{jt}\beta_{t-r_j} + w_j^{(1)}\gamma + w_j^{(2)}\lambda_{t-r_j} + \sum_{s=1}^{\mathcal{S}} d_{js}\theta_s + \xi_{jt} + \bar{\varepsilon}_{ijt}$$

where $\beta_{t-r_j}$ is a vector of preference coefficients that are parameterized to vary depending on how long the movie has been released. These parameters allow demand elasticities for time-varying movie characteristics to evolve over a movie's release window. The vector of time-invariant movie characteristics $w_j$ has been partitioned into two distinct compo-

[29]Movie ticket prices are not included as a product characteristics because they are constant across movies and over the our two year time horizon.

[30]Expected performance tiers are included in the model to control for different levels of offline WoM between movies and absorb potential endogeneity in advertising spending. Movie genres are included to allow for flexible substitution patterns across films in different genres. Further details on the role of EPT and genre in the empirical specification are contained in Section 5.1.

[31]In the econometric specification $\xi_{jt}$ serves as an econometric error term, following Berry (1994).

nents, $w_j^{(1)}$ and $w_j^{(2)}$. Movie characteristics $w_j^{(1)}$ are assumed to be valued at a constant level throughout a movie's release. Adding the time-varying coefficients, $\lambda_{t-r_j}$, for $w_j^{(2)}$ provides a flexible way to characterize the differential time paths of movie market shares as a function of their fixed characteristics. $\theta_s$ represents the consumer's utility from seeing a movie that belongs to a given performance tier $s$.

We assume that consumers consider movie genres $g$ as distinct product nests, parameterizing $\bar{\varepsilon}_{ijt}$ to allow valuations to be correlated across movies in the same genre $g$. We reserve the index $g = 0$ for the outside good. Following Cardell (1997), individual specific valuations are modeled as:

$$\bar{\varepsilon}_{ijt} = \sum_{g=0}^{\mathcal{G}} d_{jg}\zeta_{igt} + (1-\rho)\varepsilon_{ijt}$$

where $\varepsilon$ is assumed to be iid Type 1 extreme value distributed, $d_{jg}$ is a dummy variable taking the value one if movie $j$ belongs to genre $g$ and $\zeta_{igt}$ has a unique distribution such that $\bar{\varepsilon}_{ijt}$ is also iid Type 1 extreme value distributed. The 'nesting parameter' $\rho$ is restricted to lie between zero and unity and measures the degree of preference similarity between movies in the same genre.[32]

If the consumer decides not to attend a movie, $j = 0$, in period $t$ his conditional indirect utility is:

$$u_{i0t} = -\tau_t \; + \; \bar{\varepsilon}_{i0t}$$

where $\tau_t$ represents a seasonal component common to all consumers (Einav (2007)). The $\tau_t$'s are designed to pick up differences in consumer's willingness to go to theatres across weeks in a year, driven by public holidays and seasonal weather patterns. $\varepsilon_{i0t}$ represents idiosyncratic valuations specific to individual consumers.

We assume that each consumer $i$ on day $t$ chooses the movie $j$ that maximizes their utility from the set of movies that is currently available in theatres, $J_t$. To simplify notation, define the mean utility consumers receive from viewing movie $j$ in period $t$ as $\delta_{jt} = x_{jt}\beta_{t-r_j} + w_j^{(1)}\gamma + w_j^{(2)}\lambda_{t-r_j} + \sum_{s=1}^{\mathcal{S}} d_{js}\theta_s + \xi_{jt}$. With the above assumptions in place, the conditional

[32]As $\rho$ goes to 1, the within genre correlation of utilities also goes to one, and consumers perceive all movies in the same genre as perfect substitutes. When $\rho$ goes to zero, within genre correlation goes to zero and the model reduces to a simple logit demand model.

probability that consumer $i$ chooses movie $j$ on day $t$ among those showing takes the familiar nested logit form:

$$s_{ijt}(\beta,\rho) = \frac{\exp(\delta_{jt}/(1-\rho))}{\exp(I_{igt}/(1-\rho))} \frac{\exp I_{igt}}{\exp I_i t}$$

where $I_{igt}$ and $I_i$ are inclusive values defined by McFadden (1978) as:

$$I_{igt} = (1-\rho) \ln \sum_{\ell=1}^{J_g} \mathbb{1}_\ell \{\text{In Cinema}\} \exp(\delta_{\ell t}/(1-\rho))$$
$$I_{it} = \ln \left( \exp(-\tau_t) + \sum_{g=1}^{G} \exp(I_{igt}) \right)$$

and $J_g$ is the number of movies in genre $g$, such that $\sum_{g=1}^{G} J_g = \mathcal{J}$. $\mathbb{1}_\ell \{\text{In Cinema}\}$ is an indicator variable that takes the value 1 if a movie is currently showing in cinemas and zero otherwise. The addition of this 'sliding window' indicator variable accommodates the rotation of movies through the cinemas, which is an important constraint on consumer's choice (Ainslie et al. (2005)). Summing choice probabilities across consumers yields national market shares $s_{jt} = s_{ijt}$.[33]

## 5.1 Estimation & Parametrization

**Estimating Equation.** We follow Berry (1994) and invert the market shares, $s_{jt}$ to solve for mean utility, $\delta_{jt}$. This results in the familiar analytical solution for the inverted choice probabilities:

$$\begin{aligned}\ln(s_{jt}) - \ln(s_{0t}) &= \delta_{jt} - \tau_t + \rho \ln(s_{j|gt}) \\ &= x_{jt}\beta_{t-r_j} + w_j^{(1)}\gamma + w_j^{(2)}\lambda_{t-r_j} + \sum_{s=1}^{\mathcal{S}} d_{js}\theta_s + \tau_t + \rho \ln(s_{j|gt}) + \xi_{jt}\end{aligned}$$

where $s_{j|gt}$ is the share of movie $j$ within genre $g$ on day $t$. Following standard practice we treat the unobserved to the econometrician product characteristic, $\xi_{jt}$, as the error term in

[33] Appendix D outlines how we compute market shares from box office data.

the econometric model. The model is linear in parameters and the econometric error term and can be estimated using linear regression techniques. We estimate the model using data for the first three weekends of a movie's release.[34]

**Time-varying Characteristics, $x_{jt}$.** Our main goal is to identify the impact of tweets volume and tweet sentiment on the demand for movies. As outlined in Section 4.2 both evolves over time. We include pre- and post-release measures of volume in $x_{jt}$ to estimate how tweets posted at different points in time influence demand. Similarly, we include movie specific pre- and post release sentiment measures.[35] In addition to the measures of tweet volume and sentiment, we include pre- and post-release measures of advertising spending across all media. Including advertising expenditure controls for variation in demand due to firm marketing activities. For each of these characteristics we model consumer preferences to differ between the opening weekend and subsequent weekends:[36]

$$\beta_{t-r_j} = \beta_{\text{open}} \mathbb{1}\left\{0 \leq t - r_j < 3\right\} + \beta_{\text{post}} \mathbb{1}\left\{t - r_j \geq 3\right\}$$

**Time-invariant Characteristics, $w_j$.** The demand model outlined above partitions time-invariant movie characteristics into two sets. The first, $w_j^{(1)}$, contains movie characteristics for which consumers are assumed to have time-constant preferences. We include actor starpower and 2nd degree polynomials in production budget and metascore in $w_j^{(1)}$.[37] We also include the movie's CinemaScore grade, a measure of opening night consumer re-

[34]As discussed in Section 3, approximately 70 percent of a movie's earnings from it's theatrical release stem from the first three weekends of release.

[35]The two separate sentiment measures are computed from the tweets used to compute the volume measurements at a given point in time. An alternative specification would have been to include pre-release sentiment and the change between pre-release sentiment and post release sentiment on day $t$. This 'quality update' between pre- and post-release would lead to similar empirical results and looks more natural for a consumer learning framework. However the regression coefficient on the update would impose identical consumer responses to increases and decreases in sentiment.

[36]Our time-varying parameter specification is important if we want to allow pre-release measures of tweet volume, sentiment and advertising spending to have potentially stronger effects on opening weekend demand than in post-opening weekends. $\beta_{\text{post}}$ is common to weekends 2 and 3. We restrict the specification so that post-release quantities have no effect on opening weekend demand. Relaxing the assumption to allow, for example, the quantity of post-release tweets accumulated over the opening day of release to effect demand on the opening Saturday has no impact on our results.

[37]We also experimented with adding director and producer starpower following the same scoring approach as for actor starpower, and higher order polynomials in budget and metascore. None of these influenced the our quantitative findings.

views, in $w_j^{(1)}$.[38] CinemaScore grades are entered as fixed effects and used as one of our controls for *offline* WoM that impacts consumer demand independently from that posted on Twitter. We anticipate that part of the offline WoM about a movie is correlated with a movie's CinemaScore. Including these fixed effects absorbs variation in demand caused by the part of offline WoM that is correlated to CinemaScore grades.[39]

The demand model also allows for some time-invariant characteristics, $w_j^{(1)}$, to influence consumer choices differently across the release window.[40] This decrease in willingness to attend is most likely driven by declining interest & awareness.[41] The existing literature typically models declining market shares over a movie's release by estimating linear decay function that is common across movies and depends on the number of weeks that a movie has been released. Our modeling strategy relaxes this assumption in two ways. First, we flexibly model decay at the daily level using fixed effects for the number of days, rather than weeks, in release. This breaks the assumption of linear decay and means that the decline in market shares is parametrized at the same frequency as our data. Second, we allow for the 'decay fixed effects' to be a function of movie characteristics. We include separate fixed effects for movies that are part of an existing franchise and for each movie genre:[42]

[38]See Appendix A for a discussion of how CinemaScore grades are composed. The inclusion of CinemaScores in $w_j^{(1)}$ imposes that offline WoM has a constant effect on demand across a movie's release. Robustness exercises in Section 6.2 demonstrate that relaxing this assumption to allow offline WoM to have different effects across the opening weekend and post-opening period does not influence our main findings. We also demonstrate that the reduction in CinemaScore grades from the raw grades of A+, A, A-, B+, B, B-, C+, C, C-, D to A, B and "C or lower" does not influence the estimated parameters.

[39]As noted in Kuehn & Lampe (2018), CinemaScores are routinely cited by studios as a near instantaneous measure of word-of-mouth a movie will enjoy. We view offline WoM correlated with CinemaScore as potentially correlated with the volume *and* sentiment of tweets. Consumers could assign high CinemaScore grades because they view the movie as high quality or because they valued the anticipatory utility generated by pre-release buzz. The first of these correlates with tweet sentiment, and the latter with tweet volume.

[40]These time-varying parameters are designed to capture the decrease in consumer willingness to attend movies the longer they have been in theatres, partialling out the common decay patterns in consumer demand documented in Figure 1.

[41]Modelling movie decay also serves as a reduced form way to account for the 'one-time' consumption of movies which leads to a smaller market in each subsequent period.

[42]Recall that we define a movie to be part of a franchise if the plot depends on the previous movie, or is part of a larger movie franchise where story lines are not necessarily dependent on previous movies. We include 9 genres: drama, sci-fi/fantasy, action/adventure, romance, comedy, family, thriller, horror and 'other.'

$$w_j^{(2)}\lambda_{t-r_j} = \lambda_{t-r_j}^{(1)}\mathbb{1}\{\text{Is Franchise}\} + \lambda_{t-r_j}^{(2)}\mathbb{1}\{j \text{ in genre } g\}$$

The part of the decay specification for franchise movies absorbs differences in demand patterns for these movies due to higher awareness and offline WoM that this category of movies generally experiences. Franchise movies generally have a higher volume of tweets and have higher sentiment, particularly in the pre-release phase, because consumers are more aware and have higher interest in them driven, at least in part, by offline WoM. However, this is not the part of the casual effect of online WoM we want to estimate. Including franchise *times* day of release fixed effects means we are identifying online WoM effects from variation within each category on a given day, removing this source of potential bias. Intuitively this means our identifying variation comes from how differences in, for example, the volume of pre-release tweets for Spectre and the Hunger Games (which are both part of a franchise), or between Sicario and Concussion (both not part of a franchise) covary with demand on the opening Sunday, as opposed to comparing covariances between Spectre and Sicario. This within franchise variation, conditional on other product characteristics and on a given day, is the meaningful variation in pre-release volume that identifies the effect of online WoM.[43]

Genre decay patterns allow for consumers to have different willingness to attend movies of specific genres between weekends and across days within a given weekend.[44] The decision to flexibly model market share decay as a function of day of release and genre reflects the importance of correctly specifying the evolution of demand due for identifying the effects of online WoM on new product demand.

**Expected Performance Tiers (EPTs), $d_s$.** We include a movie's expected performance tier as a fixed effect to pick up consumer awareness caused by offline channels and to mitigate endogneity concerns with respect to advertising spending. EPTs are defined

[43]Franchise movies also have higher advertising budgets than non-franchise movies. Our FE's soak up this cross category difference, and identify advertising effects also from the within-day, within category variation.

[44]For example, family movies are less likely to be viewed on Fridays because the late evening show times are not convenient for families with small children, while horror and thriller movies are typically viewed earlier in their release. They also pick up differences in offline WoM that persist across movies of different genres over time.

based on pre-release expected box office earnings derived from Hollywood Stock Exchange prices.[45] Movies are partitioned into 6 mutually exclusive segments: expected earnings between $0 and $25m, $25m and $50m, $50m and $75m, $75m and $100m, $100m and $200m, and over $200m. Empirically, these fixed effects should play an important role and motivation for their inclusion mirrors that of including day of release fixed effects for franchise movies. First, movies with different expected earnings levels have different levels of consumer buzz and awareness driven by offline WoM that we want to control for. Movies that are expected to be more successful, such as *The Martian* or *Fifty Shades of Grey*, feature both higher levels of buzz and offline WoM, and a higher volume of tweets than movies such as *Chappie* or *Pompeii*. Including EPT fixed effects means that we are using variation within an EPT to identify effects, and so are using variation in tweet volume and sentiment between *The Martian* and *Fifty Shades of Grey* to identify effects rather than comparing the these movies to smaller ones. Second, advertising budgets are generally higher for movies with larger expected box office. EPT fixed effects mean that advertising elasticities are identified from covariation between ad spend and box office within a performance tier for rather than identifying variation coming from the entire cross-sectional variation. As a result we including these FEs should lead to lower, more plausible, advertising elasticities than the existing new products literature. Unlike the franchise fixed effects, we do not allow the EPT fixed effects to differentially impact consumer demand over time in the main specification.[46]

**Seasonality, $\tau$.** $\tau_t$ captures common seasonal patterns in consumers' willingness to attend theatres within a year. In our setting where variation in national demand over a movie's release helps to identify demand parameters it is also important to flexibly control for these seasonal patterns so that they are not accidentally absorbed into consumer preference parameters. We model seasonal patterns by incorporating two components of seasonal variation: (i) calendar week fixed effects, and (ii) public holiday fixed effects. Therefore $\tau_t$ is specified as:

[45] See Appendix A for further description on the Hollywood Stock Exchange, how prices are determined and how these prices translate to expected box office earnings.

[46] Robustness exercises in Section 6.2 show that allowing EPT fixed effects to differ across weekends has no influence on our results.

$$\tau_t = \sum_{c=1}^{52} \kappa_c \mathbb{1}\left\{t \text{ in calendar week } c\right\} + \sum_{d \in \{\text{Fri, Sat, Sun}\}} \omega^{(d)} \mathbb{1}\left\{t \text{ is on day } d \wedge t \text{ on Public Holiday Weekend}\right\}$$

Assuming that each weekend of the year has its own FE allows us to flexibly control for consumer demand driven by seasonal aspects such as public holidays, seasonal weather patterns and varying consumer attention to movie schedules.[47]. We control for shifts in demand across days within a public holiday weekend by allowing each calendar day within a public holiday weekend to have its own fixed effect, $\omega^{(d)}$.

## 5.2 Estimation Strategy & Identification

To estimate the demand parameters consistently we need to overcome endogeneity and omitted variable bias. These concerns are particularly relevant when estimating the effect of WoM and advertising because they are driven by strategic decisions of consumers and firms and are not randomly assigned. In our model, endogeneity concerns affect the within market share, $s_{j|gt}$, due to its' correlation with $xi_{jt}$. We adopt an IV strategy to tackle this endogeneity. Omitted variable bias is particularly important in terms of estimating online WoM because it is positively correlated to unobserved offline WoM. The rich fixed effects included in the product characteristics, using expected performance tiers, CinemaScore grades and a market share decay specification that differs between franchise and non-franchise movies and across genres, mitigates concerns about the positive correlation between tweet volume, tweet sentiment and advertising spending and offline WoM. We assume that the remaining residual variation between market shares and these variables identifies the structural parameters. The decision to adopt a rich fixed effects strategy to soak up variation stems from the lack of strong and valid instruments for each of our time-varying characteristics at the national level (Rossi (2014)). We discuss the the IV and

[47] On the latter point, consumers may generally be more interested in attending movies over Summer or Christmas weekends, not only because they are holiday periods but also because there is a 'norm' of blockbuster movies being released around these periods

fixed effect strategies in turn.

**Within Genre Shares.** The within genre market share is endogeneous because it is correlated with the unobserved characteristic, $\xi_{jt}$. We use the number of movies available in cinemas within a genre on a given day to instrument for the within genre market share, $s_{j|gt}$ (Einav (2007)). Our instrument is relevant because an increase in the number of movies within a given genre increases within genre competition, which should decrease the within genre market share of a given movie. It is excludable because the number of competing movies is not a factor that consumers likely consider when selecting what movie to watch.

**Twitter Variables.** Our approach to identifying the impact of Twitter based WoM relies on a rich set of fixed effects to control for offline WoM which would otherwise bias our estimates. These fixed effects are included to minimize the level of unexplained variation in the data, so that residual variation in market shares that is correlated with the Twitter WoM measures can be interpreted as the causal effect of Twitter on demand.[48] This fixed effects based approach to absorb endogenous sources of variation has proven successful in other contexts, and is preferred to using weak or invalid instruments which would induce their own bias (Rossi (2014), Dubé et al. (2005), Thomas (2017), Mummalaneni et al. (2019)). The threat of weak instruments is particularly relevant in our setting that relies on variation at the national level. As described in the previous section, we include four sets of fixed effects in our empirical specification to absorb the potentially endogenous variation: CinemaScore grades, expected performance tiers, franchise $\times$ day of release fixed effects and genre $\times$ day of release fixed effects. These fixed effects flexibly absorb variation between market shares and offline WoM. Conditional on the fixed effects removing the effects of offline WoM, we assume that tweet volume and sentiment are exogenous. Identification then arises from the simple timing assumption that online WoM about a movie in days prior to $t$ impact demand on day $t$.[49]

[48]Without controlling for offline WoM, correlation between market shares and Twitter would likely capture variation between demand and a measure of general WoM (i.e. online *and* offline). Measures of Twitter volume and sentiment would then serve as proxy variables for this aggregate counterpart. While this approach, under certain assumptions, would allow us to measure the impact of WoM on movie demand, the focus of this paper is to measure the causal effect online WoM on demand.

[49]Another possible cause of endogeneity is simultaneity between market shares and tweets. Recall that in Section 4.2 we defined a day $t$ as beginning at midday East Coast time and running for the next 24 hours

**Advertising Expenditure.** Advertising spending decisions may also be endogenous because the level of spending may depend endogenous quality, $\xi_{jt}$ in our model, which is observable to the movie studio but not to us as econometricians. Song et al. (2016) shows that pre-release advertising is not correlated with movie quality, suggesting endogeneity between pre-release spending and remaining unobservables is unlikely to be of first order importance. However, Song et al. (2016) and Rennhoff & Wilbur (2011) suggest that there is likely to be correlation between post-release advertising and movie quality, suggesting endogeneity concerns need to be addressed. National TV advertising comprises approximately 70 percent of movie advertising spots and 95 percent of advertising expenditure, and there is no standard and strong instrumental variable for national TV ads.[50] Thomas (2017) demonstrates that the use of fixed effects to minimize the level of unexplained variation in the data significantly reduces the endogeneity of advertising. We build on this approach by relying on our rich set of fixed effects to soak up endogeneous sources of variation. For endogeneity to bias our estimates, correlation between spending and unobserved movie quality would need to persist conditional on movie characteristics, franchise and genre-specific day of release FEs, and measurable aspects of movie quality captured by critic reviews via metascore, offline WoM via CinemaScore grades and tweet sentiment.[51] We think this is unlikely, particularly in view that our estimated advertising elasticities are substantially smaller than previously found in the new products literature.

until 11.59am on the next calendar day. This was because we wanted to capture tweets about movies from consumers who attended late sessions of a movie that spilled over into the next calendar day. Because the majority of box office receipts are earned in the late afternoon and evening we do not believe that this induces endogenity via simultaneity. However, one could be concerned that movie tickets are bought in the morning or afternoon of day $t$ and these decisions are correlated with new tweets posted over the same time period. Results using standard definitions of days, running from 00.00 until 23.59 do not influence quantitative findings below.

[50]Recent studies have attempted to estimate TV advertising effects using discontinuities across DMA borders (Shapiro (2018*b*), Shapiro (2018*a*), Sinkinson & Starc (2018), Tuchman (2018)). This approach is infeasible in our context because we have national box office data, and advertising does not differ substantially across DMAs. An alternative IV approach would be to use common cost shocks to the price of advertising to instrument for advertising expenditures (Chintagunta et al. (2006), Gordon & Hartmann (2013)). Over our time horizon there is little variation in advertising prices, and attempting to use the small amount of variation that exists presents a weak instrument problem.

[51]Similar to our construction of Twitter measures using , we measure advertising spending on day $t$ as the sum of expenditures until midday on the US East Coast.

# 6 Results

## 6.1 Average Effects

Table 2 reports the estimated parameters for the key variables in our model.[52] As discussed in Section 5.2, we instrument for $\rho$ using the number of movies in a genre. We use our rich set of fixed effects to mitigate potential endogeneity of advertising and the effects of offline WoM. The first panel of Table 2 reports the estimated coefficients for tweet volume, sentiment and advertising. We discuss how these parameters influence demand by looking at demand elasticities each of these variables in the next subsection. The market segment fixed effects, $\theta$, are intuitive: increasing in ex-ante expected box office performance and reveal substantial differences in magnitude. Interpreted through the lens of the demand model, on average consumers place higher value on attending movies with larger expected box office which leads to increased demand for these movies. We interpret this as evidence of the importance of offline buzz and awareness that is captured in the HSX stock prices. The nested logit substitution parameter, $\rho$, which reflects consumers heterogeneity in preferences across genre, is small in magnitude and estimated imprecisely. The small magnitude suggests that there is not greatly higher cross movie substitution between films that are in the same genre.[53]

Figure 3 shows estimates of the remaining fixed effects included in the demand model. These also accord with economic intuition. Panel 3a reports the differential decay patterns of movies that are part of a franchise. The results highlight that movies which are part of a franchise perform better across the opening three weekends of release, and do particularly well on the opening weekend.[54] Genre specific decay patterns are plotted in Panel 3b. Decay patterns are relatively similar across all movie genres, but with meaningful differences in levels. Controlling for other characteristics, Horror movies have stronger opening and second weekends. Family movies have similar inter week dynamics, but within a weekend

[52] OLS estimates that do not instrument for endogeneity in the within genre share are contained in Table E.1

[53] Recall that the demand model has genre $\times$ day of release fixed effects. These FE absorb most of substitution between genres into the mean utility component of the utility function. In models where decay is not specified to be genre specific, we find larger estimates of the nest substitution parameter.

[54] These results align with results in Ishihara & Moorthy (2018) which documents box office receipts are more front-loaded for sequels.

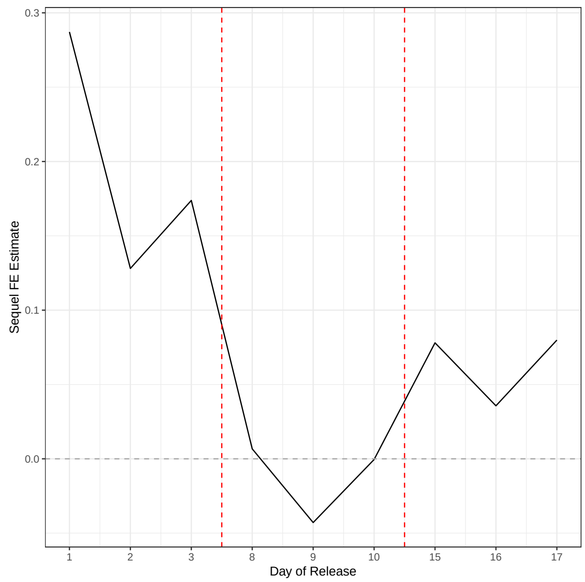
Table 2: Nested Logit Demand Estimates

| | Opening Weekend | Post Opening | Time Invariant |
|---|---|---|---|
| **Time-Varying Parameters** | | | |
| $\text{Volume}_{pre}$ | 0.301 | -0.201 | |
| | (0.135) | (0.186) | |
| $\text{Volume}_{post}$ | | 0.62 | |
| | | (0.258) | |
| $\text{Sentiment}_{pre}$ | -0.004 | -0.029 | |
| | (0.006) | (0.021) | |
| $\text{Sentiment}_{post}$ | | 0.051 | |
| | | (0.023) | |
| $\text{Ad Spend}_{pre}$ | 0.002 | 0.009 | |
| | (0.006) | (0.007) | |
| $\text{Ad Spend}_{post}$ | | 0.074 | |
| | | (0.04) | |
| **Exp. Performance Tier FE** | | | |
| $\theta_{(25,50]}$ | | | 0.935 |
| | | | (0.149) |
| $\theta_{(50,75]}$ | | | 1.385 |
| | | | (0.158) |
| $\theta_{(75,100]}$ | | | 1.742 |
| | | | (0.183) |
| $\theta_{(100,200]}$ | | | 1.92 |
| | | | (0.184) |
| $\theta_{(200,\infty]}$ | | | 2.329 |
| | | | (0.253) |
| **Nest Subst. Parameter** | | | |
| $\rho$ | | | 0.102 |
| | | | (0.066) |
| **Supplementary FE** | | | |
| CinemaScore FE | | | Yes |
| Sequel × Day of Release FE | | | Yes |
| Genre × Day of Release FE | | | Yes |
| Calendar Week FE | | | Yes |

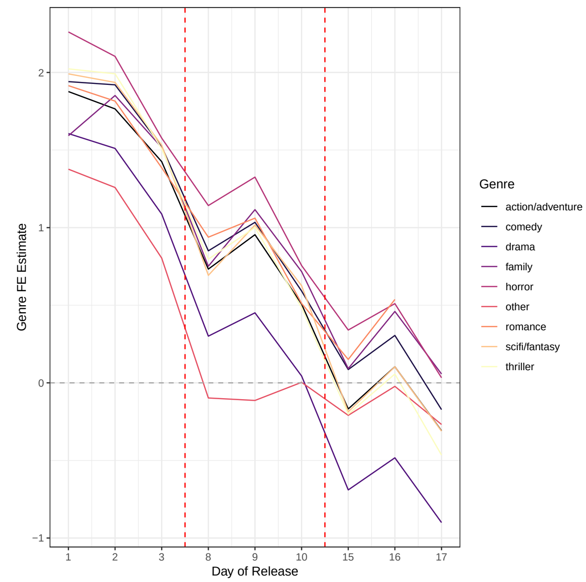
*Notes:* Key parameter estimates from IV estimation of nested logit demand model with standard errors clustered by movie reported in parentheses. As discussed in Section 5.2, we instrument for $\rho$ using the number of movies in a genre. We use our rich set of fixed effects to mitigate potential endogeneity of advertising and the effects of offline WoM. Figure 3 reports estimates of fixed effects from the model. Additional parameters for actor starpower, production budget and critic reviews are estimated but not reported.

have lower attendance on Fridays. Dramas and movies classified as 'other' are less successful and decay faster than other genres. Panel 3c highlights seasonality in demand patterns.

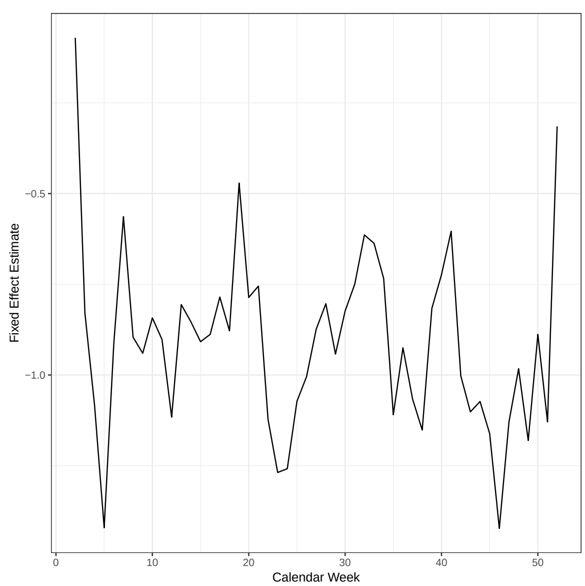
Figure 3: Fixed Effect Estimates From Demand Model



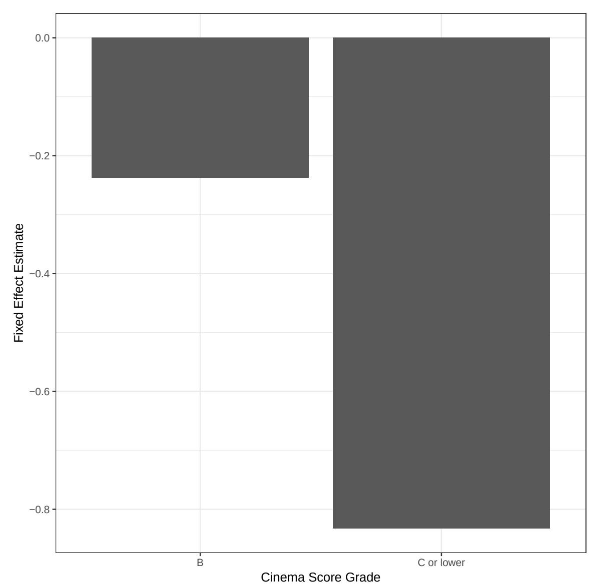

(a) Sequel FE by day




(b) Genre FE by day




(c) Calendar Week FE




(d) CinemaScore FE

*Notes*: Key Fixed Effects estimates from IV estimation of nested logit demand model outlined in Table 2. Further details of the Fixed Effect specification is outlined in Section 5.1.

Consumers are more likely to attend movies over the Christmas/New Year, the beginning and end of Summer and over the Columbus Day weekend in mid October (week 42).

Table 3: Demand Responses to Twitter WoM and Advertising

| | Own Demand Elasticities | | $\Delta s_{jt}$ w.r.t. to 0.5 Std Devn Change | |
|---|---|---|---|---|
| | Opening Weekend | Post Opening | Opening Weekend | Post Opening |
| Volume$_{pre}$ | 0.06 | -0.04 | 5.44 | -3.73 |
| Volume$_{post}$ | - | 0.08 | - | 8.10 |
| Sentiment$_{pre}$ | -0.02 | -0.17 | -1.29 | -10.41 |
| Sentiment$_{post}$ | - | 0.27 | - | 12.01 |
| Ad Spend$_{pre}$ | 0.04 | 0.18 | 1.12 | 4.67 |
| Ad Spend$_{post}$ | - | 0.12 | - | 6.58 |

*Notes:* Own demand elasticities are computed using $\hat{\eta}_{jj} = \frac{\hat{\beta}_{j,t-r_j}}{(1-\hat{\rho})} x_{jt}(1-\hat{\rho}\log(s_{j|gt})-(1-\hat{\rho})s_{jt})$ and averaged over movies. Responses to half standard deviations changes in tweet volume, sentiment and advertising expenditure is computed as $\widehat{\Delta s_{jt}} = \hat{\eta}_{jj}\frac{0.5sd(x_{jt})}{\bar{x_{jt}}}$. The magnitude of the elasticities and demand responses differ across the opening weekend and post opening periods due to the different estimated parameters for each of these phases, $\hat{\beta}_{j,open}$ and $\hat{\beta}_{j,post}$, as reported in Table 2.

These are periods where consumers have more leisure time. Panel 3d shows that movies with higher CinemaScores perform better. Combined, the fixed effect results show that the expected performance tier, CinemaScore and franchise fixed effects are absorbing important variation in demand from sources other than Twitter.

### 6.1.1 Demand Elasticities

The first two columns of Table 3 reports estimated own demand elasticities for each of the key variables of the model implied by our estimated in the top panel of Table 2. The latter two columns reports the average percentage change in daily market share due to a half standard deviation change in a variable from its mean. These results reveal that both the volume of tweets and tweet sentiment impact demand.

**Tweet Volume.** The pre-release volume elasticity is 0.055 for opening weekend demand. The elasticity maps into 5.4 percent increase in opening weekend market share if pre-release tweet volume increases by a half standard deviation. The post opening volume elasticity of 0.083 translates into an average increase demand of 8 percent in response to a half standard deviation increase in post-release tweet volume.

The results clearly demonstrate that higher Twitter volume increases demand. Increasing pre-release volume reflects larger amounts of consumer buzz and awareness that increases demand in the opening weekend. Increased tweet volume after release most likely arises from increasing awareness that stimulates demand in subsequent weekends, as well as lingering effects of buzz from consumers who have not yet seen the movie.[55]

Our volume elasticity estimates are of a similar order of magnitude to Seiler et al. (2017) and Lovett & Staelin (2016)'s estimates for TV viewing, but substantially lower than existing evidence for new product demand.[56] We find these estimates for new products implausibly large, and least for movies, and think our estimates are lower because our fixed effects strategy has separately identified the effects of online WoM from offline WoM. The fact that movies with larger ex-ante expected box office have more tweets because they are more noticeable to consumers and more heavily discussed offline is absorbed in the EPT and franchise fixed effects and not picked up by our estimated volume coefficients.

**Tweet Sentiment.** Estimates of the elasticity of opening weekend demand with respect to sentiment are small and not statistically significant. This is what one might expect, given the lack of information about a movie's true quality available pre-release. Post opening, we see larger, and statistically significant, sentiment elasticities. The estimated demand elasticity in the post opening period is 0.28 for post-release sentiment. This maps into an average market share increase of 12.2 percent in response to a quarter standard deviation increase in post release sentiment, holding pre-release sentiment fixed. Similar to the pre-release tweet volume, the post opening demand elasticity with respect to pre-release sentiment is negative. This means an increase in pre-release sentiment for a given level of post-release sentiment lowers demand.

These two post release sentiment elasticities are consistent with a mechanism of con-

[55]It is also important to note that the elasticity of pre-release volume in post opening phase is negative. If a given level of post-release tweets is associated with a higher level of pre-release tweets there has been a decrease in the relative amount of buzz or awareness about a movie after release. This then leads to a contraction in demand.

[56]Seiler et al. (2017) reports an estimated volume elasticity of 0.016 for well established TV series in China. Lovett & Staelin (2016) document an of elasticity 0.04 for the introduction of a new series based on an existing comic book. Recent studies on the movie industry find estimates of online volume elasticities of 0.57 and 0.64 (Liu (2006), Kim & Hanssens (2017)) and estimates for new product demand generally lie between 0.6 and 1.

sumers learning about movie quality. When post-release sentiment increases, consumers infer that the movie is of higher quality and respond by going to see that movie in cinemas. Similarly, when pre-release sentiment is high, but post-release sentiment does not match this enthusiasm, consumers lower their quality expectations and are less likely to go to the movie.

**Advertising Expenditure.** Estimated pre- and post-release advertising elasticities are much lower than those typically reported in previous studies. For example, Gopinath et al. (2013) find a pre-release ad spending elasticity of 0.46 and a post release elasticity of 0.52. This discrepancy is because our advertising elasticities are estimated from variation within a market segment and franchise category. This means that the segment and franchise fixed effects are capturing a large part of potential awareness & quality signaling that advertising spending traditionally captures, leading to estimated ad spending elasticities that are more reasonable in magnitude. For example, the opening weekend demand elasticity for pre-release advertising of 0.04 is similar in magnitude to recent work estimating demand elasticities to TV advertising (Gordon & Hartmann (2013), Shapiro (2018*b*), Tuchman (2018)). Similarly, our post opening elasticities are similar to recent work studying the impact of movie advertising on consumer interest (Kim & Hanssens (2017)).

## 6.2 Validity of Empirical Strategy

Our decision to adopt a rich fixed effects to address the potential bias in our estimates of online WoM due to offline WoM and endogenous advertising could still be biased. To address this concern, this section demonstrates the stability of our reported parameter estimates to alternative empirical specifications.

**Movie Fixed Effects.** Estimates replacing segment fixed effects and other time-invariant movie characteristics with movie fixed effects are reported in Table E.2. Movie fixed effects may be able to better absorb potential variation driven by unobserved characteristics, but comes at the expense of being unable to estimate coefficients for pre-release measures because they do not vary over time. The results show that the parameters that could be estimated in both models are similar, with estimated coefficients in each model lying inside the standard error bands of the others.[57] This provides strong support for the fixed effects strategy that we have implemented, showing that our post release parameters are not biased relative to the movie fixed effects estimates *and* they allow us to estimate demand elasticities for pre-release tweet volume, sentiment and advertising spending.[58]

**Evolving Offline WoM.** Our empirical strategy uses expected performance tier and CinemaScore fixed effects as controls for offline WoM. We parameterize these FEs to be constant over the first three weekends of release. As a result, parameter estimates for tweet volume and sentiment could be picking up the part offline WoM that is evolving over a movie's release. Tables E.4, E.5 and E.6 in Appendix E show that this is unlikely to be the case. Across these tables we allow the EPT and CinemaScore fixed effects to be differ between the opening and weekend and post-opening weekends, adopting the same structure as the parameters for tweet volume and sentiment. Across each specification the estimated demand parameters remain quantitatively similar.[59]

**Residual Offline WoM.** The rich fixed effects strategy we have adopted aims to absorb offline WoM so that the demand elasticities for tweet volume and sentiment reported in Table 3 measure the causal effect of online WoM. The causal interpretation of these elasticities rests on the assumption we have picked up as much offline WoM in our fixed effects structure as possible so that there is no remaining unmodeled correlation between $\xi_{jt}$ and the measurements of tweet volume and/or sentiment. We have two sets of supplementary results to demonstrate that this is likely to be the case: First, we have re-estimated the model using less coarse measures of the CinemaScore to soak up additional variation in the data that could be due to offline WoM. Using the raw CinemaScore grades movies yielded similar results (Tables E.7) .[60] Second, we estimated models including the additional data

[57]We also tried alternative definitions and numbers of movie segments to better approximate the movie fixed effects. Changes in the thresholds of the six segments had little impact on estimated coefficients, nor did increasing the number of segments to eight.

[58]Recall that our fixed effect strategy uses expected performance tiers, CinemaScore grades and modelling decay at using daily FEs for each genre and franchise category.

[59]We have also tried allowing movie segment and cinemascore FE to to differ at the weekends since release. These specifications do not alter our findings.

[60]We also and an alternative grouping that separated 'A+' from other A rated moves in Table E.8, yielding similar estimates. We chose to separate out A+ as its own FE because movies in this category had different box office dynamics.

from the HSX. The idea here is that the HSX market has more information about $\xi_{jt}$ than we do as econometricians, so that controlling for it removes potential bias. We estimated specifications with the raw pre-release prediction[61], the difference between the pre-release prediction and the updated prediction at the Thursday close per weekend[62], and the sign of the difference between the pre-release prediction and the updated prediction at the Thursday close per weekend (Tables E.9, E.10 and E.11 respectively). Across all specifications the results remain similar.

## 6.3 The Differential Effects of Online WoM Across Movies

The results presented above assume that the impact of Twitter on demand is common across all movies. This masks potential heterogeneity not only in the magnitudes of the effects, but also the mechanisms through which Twitter impacts demand across different types of movies. We investigate these differences by allowing the time-varying parameters, $\beta_{t-r_j}$, to differ across two important characteristics: (i) whether a movie is part of an existing franchise, and (ii) expected performance tiers.

### 6.3.1 Franchise Heterogeneity

Table 4 reports own demand elasticities that are parameterized to differ between franchise and non-franchise movies.[63] The top panel reports elasticities for movies that are part of a franchise whilst the lower panel reports results for stand alone movies. The results reveal substantial differences in the magnitude of the elasticities across movie types, revealing different mechanisms through influencing demand.

Looking first at the estimates for movies that are part of a franchise, we see that pre-release volume elasticity, 0.17, is more than 3 times as large as the average elasticity reported in Table 3. The results also suggest that, for franchises, the effect of pre-release

[61] This is designed to capture differences in the levels of in $\xi_j$.

[62] This is designed to capture changes in $\xi_{jt}$ over time. We compute the change in HSX prediction as the pre-release prediction minus HSX stock price on the Thursday close. The Thursday close was chosen because we didn't want Box Office performance of the current weekend to be included in it's calculation. We allowed the coefficient on the change to be different for positive and negative changes.

[63] We re-estimate the demand model parameterized in Section 5.1, allowing time-varying parameters, $\beta$ to differ for franchise and non-franchise movies. The estimated parameters that underly these demand elasticities are reported in Table F.1 in Appendix F.

Table 4: Own Demand Elasticities with Franchise Heterogeneity

| | Own Demand Elasticities | |
|---|---|---|
| | Opening Weekend | Post Opening |
| **Franchise = True** | | |
| $\text{Volume}_{pre}$ | 0.17 | 0.08 |
| $\text{Volume}_{post}$ | - | 0.05 |
| $\text{Sentiment}_{pre}$ | 0.05 | 0.10 |
| $\text{Sentiment}_{post}$ | - | 0.00 |
| **Franchise = False** | | |
| $\text{Volume}_{pre}$ | 0.02 | -0.16 |
| $\text{Volume}_{post}$ | - | 0.23 |
| $\text{Sentiment}_{pre}$ | -0.07 | -0.52 |
| $\text{Sentiment}_{post}$ | - | 0.71 |

*Notes:* Own demand elasticities are computed using the formula's outlines in Table 3.

volume spills over into the post-opening period. The elasticity of post opening demand with respect to pre-release volume is 0.078, although it is estimated imprecisely. These results highlight the important role of pre-release consumer buzz in generating demand for franchise films, even after the opening weekend.[64]

The estimates also reveal that demand responds to pre-release sentiment, with an opening weekend elasticity of 0.05 which grows in magnitude to 0.1 in the weekends after release. This effect is different than for the average movie, where there was no effect of pre-release sentiment. Our results suggest that the amount of 'good news' in pre-release discussion is an important component of pre-release consumer buzz for franchise films, in addition to the volume of tweets. Here, an increase in the positivity in pre-release further heightens consumer anticipation. The heightened anticipation triggers increases in demand over both phases of release.[65] By contrast, post-release sentiment does *not* influence demand, suggesting that social learning and information diffusion in the post-release phase does not influence demand for franchise movies. In sum, our results indicate that online WoM is only

[64]We interpret the mechanism that generates this elasticity to be predominately pre-release buzz, rather than awareness, because consumers generally know that these movies are being released.

[65]An alternative explanation is that the quality these movies is revealed in pre-release sentiment. This explanation depends on consumers having accurate information about movie quality pre-release, for example by using a combination of previous movies' quality and the content of the pre-release advertising.

important for generating buzz and awareness for franchise movies, and not for revealing a movie's true quality.

The results in the lower half of Table 4 reveal different mechanisms at work for movies that are not part of an existing franchise. We defer discussion of these results to the next section, which provides a clearer picture of these patterns.

#### 6.3.2 Heterogeneity across Expected Performance Tiers

We now turn to investigate differences in effect sizes across another important dimension of movie heterogeneity - ex-ante expected performance. To simplify exposition, the six market segments are mapped into three categories - small, medium and large to interact with tweet volume, sentiment and advertising spending.[66] Table 5 reports own demand elasticities.[67] Each panel reports estimates for a given category, revealing significant heterogeneity in the impact of online WoM across movies with different ex-ante expected performance.

The results reveal large and positive elasticities for the impact of pre-release volume and sentiment on opening weekend demand for large movies. The large movies in our sample are all blockbuster franchises, such as *Jurassic World*, *Avengers: Age of Ultron* and *The Hobbit: Battle of Five Armies*. The point estimates for pre-release volume and sentiment, 0.34 and 0.08 respectively, are larger than documented in the previous section because the large franchises are where the effect of pre-release buzz is greatest. The large and opposing coefficients for the impact of sentiment over the post-opening phase, 0.93 and -0.81, is driven by near perfect correlation in between the pre- and post-release sentiment scores among this category.[68]

Looking across the medium and small tier movies we see similarity to the elasticities for non-franchise movies. The demand elasticity for mid-tier movies on the volume of tweets

[66] The 'small' category includes all movies with expected opening month box office between 0 and 74.99 million USD. The 'medium' category includes movies with expected opening month box office between 75 and 200 million USD. Movies in the large category have expected opening month box office over 200 million USD. Expected opening month box office was computed from the average HSX price over the 7 days immediately prior to a movie's release. We still include the six initial segments as fixed effects in the demand model. This reduction to three categories simplifies the discussion of the estimates without changing the qualitative results.

[67] The estimated parameters that underly these demand elasticities are reported in Table F.2 in Appendix F.

[68] The Pearson correlation between them is 0.96.

Table 5: Demand Responses to Twitter WoM and Advertising with Segment Heterogeneity

| | Own Demand Elasticities | |
|---|---|---|
| | Opening Weekend | Post Opening |
| **Exp. Performance Tier = Large** | | |
| $\text{Volume}_{pre}$ | 0.34 | 0.19 |
| $\text{Volume}_{post}$ | - | 0.11 |
| $\text{Sentiment}_{pre}$ | 0.09 | 0.93 |
| $\text{Sentiment}_{post}$ | - | -0.81 |
| **Exp. Performance Tier = Medium** | | |
| $\text{Volume}_{pre}$ | 0.01 | -0.17 |
| $\text{Volume}_{post}$ | - | 0.23 |
| $\text{Sentiment}_{pre}$ | 0.00 | 0.14 |
| $\text{Sentiment}_{post}$ | - | -0.20 |
| **Exp. Performance Tier = Small** | | |
| $\text{Volume}_{pre}$ | 0.10 | 0.10 |
| $\text{Volume}_{post}$ | - | 0.07 |
| $\text{Sentiment}_{pre}$ | -0.06 | -0.52 |
| $\text{Sentiment}_{post}$ | - | 0.70 |

*Notes:* Own demand elasticities are computed using the formula's outlines in Table 3.

posted after a movie has been released, 0.23, is identical to the non-franchise estimate reported in Table 4 Our results provides evidence towards the mechanism of expanding awareness in the post-opening phase among mid-tier movies. Mid-tier movies are partially crowded out of consumer's attention over the opening weekend by the large volume of pre-release tweets about franchise movies. After the opening weekend Twitter users become increasingly aware of the subset of these mid-tier movies with high post-release volume by seeing tweets appear in their feed after the opening weekend. The expanded awareness translates into more consumers seeing the movie in the weeks after a movies initial release.[69] In contrast to mid-tier movies, information diffusion about movie quality captured in tweet sentiment is concentrated towards movies with lower pre-release earnings expectations. The post-release sentiment elasticity of 0.7 for small movies matches the non-franchise movie estimate. These smaller movies are the ones with the highest uncertainty about

[69]Our post-release volume elasticities provide evidence against a social consumption motive. If social consumption is the main mechanism driving consumption we would expect to see large post-release volume elasticities across all movie segments.

movie quality. Post-release sentiment expressed on Twitter provides an important quality gauge where consumers respond to the positive experiences expressed by others, attending the movie in the following weeks. This large demand response is also driven by the fact that Twitter users see more movies than non-Twitter users and are therefore likely to view smaller movies. When they decide to what to see, they choose based on the quality perceptions they read online.

# 7 Concluding Remarks

Understanding the impact of online WoM on consumers' purchases of new products in increasingly important with the growing prevalence of sites such as Twitter, which provide easy access to decision relevant information, in consumers' daily lives. Consumer demand can be driven by two important metrics: (i) the volume of posts about a product because it increases awareness and product buzz, and (ii) the sentiment expressed in posts which provides an avenue for information diffusion via social learning. In this paper we estimated the impact of these two metrics on the demand for wide release movies over their first three weeks in theatres. A key ingredient in the analysis is granular data on the volume and sentiment of Twitter posts for each movie released over a two year horizon. The Twitter data is integrated with a comprehensive movie data to provide a complete overview of the industry. We used the data to construct a nested logit model of consumer demand for movies that allows the volume and sentiment elasticities to differ across a between a movie's opening weekend and post-opening weekends, capturing the differential importance of buzz, awareness and sentiment across a movie's release window. An important ingredient in the demand model is a rich set of fixed effects to absorb WoM from external sources allowing us to isolate the online WoM channels.

Two novel findings emerge from our analysis. First, we show that both the awareness & buzz as well as the information diffusion channel are important determinants of demand for movies, although our estimated demand elasticities are substantially lower than existing estimates for new products. Across all movie, our estimated demand elasticities for the volume of tweets, 0.05 for pre-release volume on opening weekend demand, and 0.08 for post-release volume on post-opening demand. These estimated elasticities are an order of magnitude lower than existing estimates for new product demand that generally range from 0.6 to above 1. We also show that demand responds to post-release sentiment. For a fixed level of pre-release sentiment we show that increases in post-release sentiment, which are an positive update in consumer perceptions of movie quality, increase demand with an elasticity of 0.27. Similar to our volume estimates, our sentiment elasticity is substantially lower than existing work on new product demand that typically find estimates between 0.4 and 0.5. Our lower estimated elasticities stem from the rich set of fixed effects included in our demand model which soaked up endogenous variation between online WoM and unobserved external sources of information such as offline WoM.

Furthermore, our analysis demonstrates that the mechanisms through which online WoM influences demand differ across movies with different attributes. The results highlight that pre-release volume is most important for movies that are part of an existing franchise, with a demand elasticity of 0.17 on the release weekend. The responsiveness of opening weekend demand to pre-release volume is even larger, with an elasticity of 0.34 for 'blockbuster movies' which are franchise movies that have high ex-ante predictions of box office earnings. These results emphasize that pre-release consumer buzz is a key factor driving theatrical attendance of consumers for these types of movies. In contrast, we find that demand for movies with lower expected box office is primarily influenced by the sentiment expressed in tweets by consumers who have already seen the movie. For these movies, online WoM facilitates information transmission about perceived movie quality among peers. The demand elasticity for post-release sentiment is 0.7, highlights the large shift in demand in response to increased expected movie quality. Demand for movies with 'mid-tier' expected success is sensitive to the volume of tweets in the post-release phase, suggesting maintaining and expanding consumer awareness after release is important for sustained success.

**Managerial Implications.** Our findings have direct implications for marketing practice for all new products that rely on online WoM to generate demand. First, the estimated demand elasticities reveal the magnitude of these impact of online WOM on demand is smaller than previously thought. This demonstrates that online WoM may not be the panacea that existing work suggests. This is important for two reasons: (i) because online

WoM is currently perceived to be the most effective forms of marketing it has shifting marketers attention away from traditional marketing practice (WOMMA 2013), and (ii) the difficulty industry experienced connecting business outcomes to social media may, in part, be driven by the real effect sized being much smaller than anticipated (American Marketing Association 2013).

Our results also suggest that firms should not rely on a 'one size fits all' approach to understanding & harnessing the impact of online WoM around a product's release. For new products with large ex-ante expected sales or for brand extensions via new versions with similar fundamental features to a firm's existing product line, marketing activities should focus on increasing the online WoM before a product's release to generate buzz. Examples of products where stimulating pre-release volume could be important include, but are not limited to, new models of established smartphones, sequels of videogames and newly released music from established artists. Seeding this WoM directly through firms starting the conversation or via influencers by has been shown to increase demand (Lovett & Staelin (2016), Gong et al. (2018)). Visible and easy to use hashtags that can be added add to tweets are also a key ingredient so that it is easy for consumers access the volume of posts and tap into the pre-release buzz. On the other hand, for products with lower expected sales or that feature a high degree of quality uncertainty pre-release our findings suggest online WoM that generates increased positive sentiment after the initial launch is important for stimulating demand. Example marketing strategies to enhance post-release sentiment include encouraging & rewarding consumers to leave reviews, starting and/or engaging in conversations focused on reviewing a product, or building social media advertising campaigns using existing positive reviews as content.

**Future Directions.** Our work opens several avenues for further research into understanding the impact of Twitter and online WoM more generally on new product markets. As a next step one could create an algorithm to isolate consumer buzz allowing separation of the buzz and awareness mechanisms that we jointly measure with the volume of tweets. An alternative avenue would be to adopt Deep Learning techniques to identify other salient aspects of the online WoM not captured by the volume and sentiment metrics adopted in this paper (Liu et al. (2018)). Another interesting avenue for further research would consider the interaction between online WoM and other types of WoM such as offline or via private messages online which we typically don't observe (Schweidel & Moe (2014)). Alternatively, one might attempt to quantify the interaction between traditional advertising strategies and WoM, investigating how traditional marketing practices can influence the volume and sentiment of consumer WoM (Campbell et al. (2017)). All of these avenues are left for future research.

# A Additional Data Sources

*Advertising Data.* Advertising spending is collected from Nielsen Ad Intel. The Ad Intel database covers advertising across TV, Radio, Internet, Newspaper & Magazines, Cinema, Billboards and Coupons. We extract advertising spending data from the years 2013 to 2016 for all wide release movies. Where relevant, nationwide spending is combined with local spending data at the DMA level. Each movies total advertising spend is broken into two parts: (i) pre-release advertising spend - which measures the amount of pre-release spending in the month before a movie is released, and (ii) post-release spending - measuring all ad spend from the release day until the end of the modelling period.

*Movie Characteristics.* We also collect information on a movie's title, genre, MPAA rating, release date, production budget, and leading actors from Box Office Mojo. In the event that the production budget is not available on Box Office Mojo, we use two additional data sources to supplement the missing data. First we use budget estimates from the website "The Numbers" - a leading free provider of movie data. If the budget is not available on either Box Office Mojo or The Numbers, we use information from Wikipedia.[70] We construct a dummy variable "series" that takes the value of one if the movie is part of a larger series, and zero otherwise.[71]

*Actor Starpower.* We use the lead actors listed on Box Office Mojo combined with data on Oscar award nominations and winners to construct an index of actor starpower. Our starpower measure assigns to each actor a score that sums their number of nominations and wins of Oscars. Our measure for starpower for an actor, $a$, starring in a movie in year $t$ is then $\text{Actor Starpower}_{at} = \sum_{\tau<t} \text{Nominations}_a + 2 \times \sum_{\tau<t} \text{Awards}_a$. The starpower measure for a movie, $j$ is then the sum of these scores across lead actors, $\text{starpower}_{jt} = \sum_{a \in \mathcal{A}} \text{Actor Starpower}_{at}$.

*Critic Reviews.* These movie characteristics are complemented with two measures of movie quality. Our first measure of movie quality is the "metascore" obtained from Metacritic - an review aggregator for media products. Metascore is a weighted average of reviews

[70] For the potential 222 Friday released movies in our data, we obtained estimates for all but one movie - Disney's Monkey Kingdom.

[71] Distinguishing between series, as defined here, and "sequel" where story lines across movies are interdependent does not change our empirical results.

from respected movie critics, with the weights determined by a critic's fame, stature, and volume of reviews. The output of this weighted average is a movie specific score ranging from 0 to 100.[72]

*Offline Consumer Reviews.* We complement the metascore - which is a critic based measure of movie quality, with a movie's Cinemascore - a consumer based measure. Critic and consumer scores frequent differ because they capture alternative aspects of film quality. A movie's Cinemascore is a measure of quality based on nationwide opening night polling at movie theatres.[73] Selected opening night audience members at cinemas USA fill out ballot cards at the conclusion of a movie, assigning letter grades from A through to F. These individual scores are then aggregated to assign each movie a grade from A+ to F. Cinemascore do not poll audiences for all wide release movies on a given weekend, and tend to omit films that are forecast to perform badly.

*Hollywood Stock Exchange.* The Hollywood Stock Exchange is an online prediction market where market participants can trade virtual movie stocks. Each movie that will be released in the US is listed for 'initial public offering' on the HSX when a movie's development is publicized. Once a movie's stock is listed , and trading at market prices that reflect the market's collective expectation of the movie's first four weekend revenues as a wide release. For example, if a particular movie stock trades at "H$60.00", the market is predicting that the movie will gross US$60 million at the box office in the first four weekends of wide release. Similar to real life stock markets, new information about a movie emerges over time that can influence trader beliefs. Participants who believe in a box office success is likely higher than the current market price reflect buy the stock, whereas those who believe the opposite can sell it. Spann & Skiera (2003) and Foutz & Jank (2010) demonstrate that HSX trading prices close to the release date provide reliable

[72] Movie critics do not always provide their review on the range from 0-100. Metacritic uses a consistent set of grade conversion scales to map letter-graded and 4-star graded reviews into the $[0, 100]$ interval before constructing the weighted average. Movies with metascores above 60 are considered as favorably reviewed, scores between between 40-60 are considered mixed, and movies receiving below 40 are considered to have unfavorable reviews.

[73] We have chosen to use a movie's Cinemascore over a consumer reviews posted on review websites like Rotten Tomatoes or IMDb because the Cinemascore is a static measure. That is, the Cinemascore does not evolve over time. Using alternatives like IMDb or Rotten Tomatoes often means using a score that is a combination consumer reviews obtained over the theatrical release window and DVD release window, particularly when accessing these scores retrospectively.

forecasts of predicted box office performance. Our analysis leverages these results, using the average closing price over the seven days prior to a movie's release as a measure of predicted box office performance. We use these predictions to construct 'market segments', to use as fixed effects designed to absorb aspects of consumer demand, such as general consumer awareness, that are common within each segment. Further details of the market segment construction are included in Section 5.1.

# B The VADER Sentiment Lexicon

The VADER lexicon consists of approximately 7500 lexical features classified as either positive or negative. These features include a list of words, each classified as either positive or negative, and sentiment laden emoticons and emojis. This binary classification of words, emoticons and emojis as positive or negative is complemented with a measure of sentiment intensity for each feature of the lexicon. The standard way to classify text using a sentiment lexicon is to sum the valence scores of features in a text and compare the score to a set of thresholds. The position of the valence score relative to these thresholds then determines whether a tweet is classified as positive, negative or neutral.[74]

VADER extends this standard approach by incorporating 5 heuristics based on grammatical and syntactical cues to convey changes to sentiment intensity: (1) Punctuation - exclamation points (!) increase the valence of the preceding word without changing its semantic orientation. For example "The movie was great!" is scored as more positive than "The movie was great" (2) Capitalization - an ALL-CAPS written sentiment laden word increases sentiment intensity. For example, "<movie name> was the BEST movie of the year" receives a more positive score than "<movie name> was the best movie of the year." (3) Degree Intensifiers - words that describe intensity in language also modify the intensity of the sentiment. "The movie was extremely good" is scored higher than "the movie was good" which, in turn, is scored higher than "the movie was marginally good." (4) Contrastive Conjunctions - words like 'but' indicate a shift in sentiment polarity. Words that appear after the conjunction are dominant in determining polarity. For example "<actor name> gave a strong performance, but the movie was still horrible" has mixed sentiment but is scored as negative. (5) Trigram negation - "the movie wasn't all that great" is scored negatively. This is because "wasn't" negates the sentiment score of "great" because the lexicon examines the trigram before a sentiment laden word.

[74] In sentiment lexicons that only classify words as positive or negative the approach is different. They look at the proportion of positive and negative words in a piece of text and use these to classify a tweet as positive or negative.

# C Model Free Evidence

We present some model-free results linking Twitter variables to the box office performance of movies. In particular we investigate how pre-release tweet volume and sentiment relate to opening night box office performance, and how of post-release measures of the same variables impact the the evolution of box office receipts over the first three weekends of theatrical release. These results motivate aspects of the structural demand model we specify in Section 5.

Figure 2 in Section 4.2 demonstrated tweeting about a movie begins well before theatrical release with a large degree of heterogeneity across different movies. To unpack how these differences correlate with opening night revenues, we run OLS regressions of the log of opening night box office on the volume of pre-release tweets and our measure of pre-release sentiment. The results are reported in Table C.1.

Table C.1: OLS Estimates - Opening Night Box Office Revenue

| | **Dep. Var: log(box office)** | | | | |
|---|---|---|---|---|---|
| | **Model 1** | **Model 2** | **Model 3** | **Model 4** | **Model 5** |
| log(tweet volume) | 0.59*** | 0.56*** | 0.52*** | 0.40*** | 0.39*** |
| | (0.05) | (0.05) | (0.06) | (0.05) | (0.05) |
| log(pos-neg ratio) | 0.05 | 0.06 | 0.02 | 0.01 | 0.01 |
| | (0.08) | (0.07) | (0.07) | (0.07) | (0.06) |
| log(prerel adspend) | | 0.21 | 0.19 | 0.08 | 0.06 |
| | | (0.11) | (0.11) | (0.06) | (0.06) |
| Metascore | No | No | Yes | Yes | Yes |
| Movie Characteristics | No | No | No | Yes | Yes |
| Month Controls | No | No | No | No | Yes |
| N | 161 | 159 | 159 | 158 | 158 |
| Adj. R-squared | 0.49 | 0.52 | 0.54 | 0.67 | 0.70 |

*Notes:* Parameter estimates from OLS regression of opening night box office revenue (in millions of USD) on Twitter variables, advertising spending and controls. HC1 robust standard errors in parentheses.

Columns (1) - (5) of Table C.1 show variants of this model with different sets of conditioning variables. Column (1) reports results that only include these variables in the regression specification. The results show that the volume of tweets is positive related to opening night box office, whereas pre-release sentiment is neither statistically or economically significant. The qualitative magnitude of estimated coefficient on pre-release volume is unaffected by the inclusion of pre-release advertising spending (Column (2)), critic reviews (Column (3)), movie characteristics (Column (4)) and release month controls (Column (5)).[75] Our results suggest pre-release tweet volume captures an additional force over and above studio marketing and movie quality, and is suggestive of a mechanism where tweets represent pre-release buzz and generate awareness. The fact that pre-release sentiment is not correlated with opening night returns suggests the information content of pre-release tweets does not reveal a dimension of quality consumers use in their decision making. This is makes intuitive sense because Twitter users cannot have not seen the movie, and therefore their tweets cannot reveal information about consumer experiences of movie quality.[76]

We now turn to the relationship between post release Twitter activity and evolution of box office. To study this relationship we consider how box office revenue of each movie declines over time relative to it's opening night earnings. [77] Thus box office decay for movie $j$ after $t$ days in theatres is defined as the ratio of box office earnings on day $t$ compared to opening night ($t = 0$), $\text{decay}_{jt} = \text{box office}_{jt}/\text{box office}_{j0}$. This measure of decay is then regressed against post-release measures of tweet volume and sentiment along with a detailed set of control variables.[78] Regressions are run at a daily level, over the opening three weekends of a movie's release allowing us to flexibly capture how the relationship

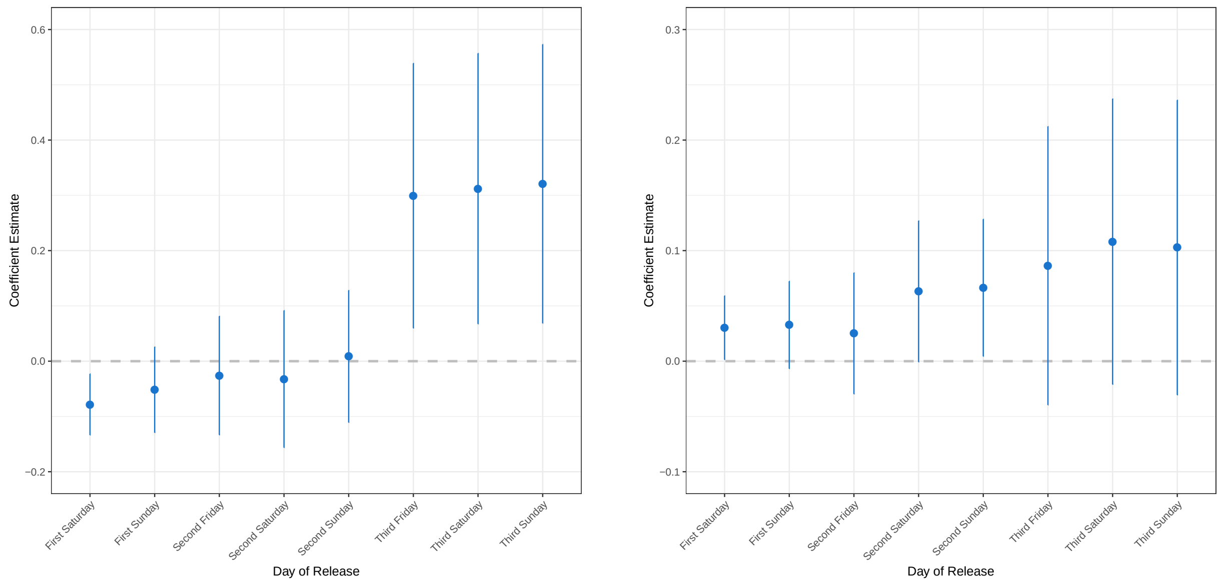
[75]Opening night box office regressions do not include CinemaScore as a control variable. This is because a movie's CinemaScore comes from polls conducted on the opening night at cinema complexes at the conclusion of a showing. Thus, they cannot have a direct impact on movie performance.

[76]A more detailed exploration of pre-release tweeting behaviour reveals that this indeed the case. Pre-release tweets generally contain messages of anticipated excitement of viewing the movie when it is released, or are retweeting the movie's trailer, official poster or other image.

[77]In addition to these regression specifications using 'decay' as the dependent variable, we have run identical regressions using log box office as the dependent variable. Table modelfree_box in Appendix C.1 reports the the results of these regressions. The main differences in the estimated coefficients are that (i) there is no relationship between post release sentiment and daily box office earnings, and (ii) post-release Twitter volume has a positive and statistically significant relationship with box office on each day.

[78]Controls included in the regression specifications are pre-release tweet volume and sentiment, pre- and post-release advertising spending, CinemaScore ratings to capture offline WoM, critic reviews via metascore, production budget, actor starpower, a dummy variable for movies that are a part of a series/sequel, fixed effects for month of release and a dummy variable for whether day $t$ falls on a long weekend. For all continuous variables we include its log-ged value as the regressor, so that estimated coefficients can be interpreted as elasticities. Post-release advertising spending on day $t$ is measures as the sum of advertising expenditures from the day of a movie's release until day $t - 1$.

Figure C.1: Impact of Post Release Twitter Measures on Box Office Decay




(a) Post release Volume Coefficients

(b) Post release Sentiment Coefficients

*Note*: Figure reports coefficient estimates from a regression of box office decay against the volume of tweets in the post-release period and sentiment measured from them. Decay for movie $j$ after $t$ days in release is defined as $\text{decay}_{jt} = \text{box office}_{jt}/\text{box office}_{j0}$. Regressions are estimated for each day separately. Controls included in the regression specifications are pre-release tweet volume and sentiment, pre- and post-release advertising spending, CinemaScore ratings to capture offline WoM, critic reviews via metascore, production budget, actor star-power, a dummy variable for movies that are a part of a series/sequel, fixed effects for month of release and a dummy variable for whether day $t$ falls on a long weekend. For all continuous variables we include its log-ged value as the regressor, so that estimated coefficients can be interpreted as elasticities. Post-release advertising spending on day $t$ is measures as the sum of advertising expenditures from the day of a movie's release until day $t-1$. 95% confidence intervals, represented by the solid lines are computed using HC1 heteroskedasticity robust standard errors.

develops over time. Figure C.1 plots the estimated coefficient (as a solid dot) and 95 percent confidence interval (as a line).[79] Panel (a) reports the estimated daily coefficients for post release tweet volume and Panel (b) reports the estimated coefficients for sentiment.

The coefficient plots reveal positive relationships between movie decay and post release Twitter measures that increase over time. Turning first to post release tweet volume, Panel (a) reveals the development of a positive relationship between post release tweet volume

[79] 95% confidence intervals are computed using HC1 heteroskedasticity robust standard errors. Table C.1.1 in Appendix C.1 reports the full set of regression coefficients for each of the 8 regressions, one for each day.

and our decay measure into the third weekend of release. This means that highly tweeted about movies in the post-release period have stronger performance for longer. The results also show that post-release sentiment is positively associated with our measure of decay. Panel (b) highlights the positive association between box office decay and post release sentiment expressed on Twitter. This increase in the estimated coefficient over time is consistent with a mechanism of information diffusion about movie quality driven by social learning. Because the regressions underlying these plots have conditioned on advertising, critic reviews and offline WoM (via CinemaScore), the results provide suggestive evidence of a relationship between tweet sentiment and the success of movies.

## C.1 Supplementary Regression Tables

### C.1.1 Evolution of Box Office Decay

The estimating equation for the box office decay of movie $j$ on day $t$ take the form:

$$
\begin{aligned}
log(\text{decay})_{jt} = \beta_0 &+ \beta_{11} log(\text{pre-release volume})_{jt} \beta_{12} log(\text{post-release volume})_{jt} \\
&+ \beta_2 log(\text{post-release sentiment})_{jt} + \beta_{31} log(\text{pre-release ad spend})_{jt} \\
&+ \beta_{32} log(\text{post-release ad spend})_{jt} + \beta_4 log(\text{metascore})_j + \beta_5 \text{consumer review}_j \\
&+ X_j \gamma + Z_t \delta + \varepsilon_{jt}
\end{aligned}
$$

The results are reported in Table C.2 Columns (1) to (8) show the regression coefficients estimated for each day, beginning the second day of release (i.e the First Saturday), using our preferred specification that includes all conditioning variables.

### C.1.2 Evolution of Box Office Revenue

The estimating equation for the box office earnings of movie $j$ on day $t$ take the form:

$$
\begin{aligned}
log(\text{box office})_{jt} = \beta_0 &+ \beta_{11} log(\text{pre-release volume})_{jt} + \beta_{12} log(\text{post-release volume})_{jt} \\
&+ \beta_2 log(\text{post-release sentiment})_{jt} \\
&+ \beta_{31} log(\text{pre-release ad spend})_{jt} + \beta_{32} log(\text{post-release ad spend})_{jt} \\
&+ \beta_4 log(\text{metascore})_j + \beta_5 \text{consumer review}_j + X_j\gamma + Z_t\delta + \varepsilon_{jt}
\end{aligned}
$$

The results are reported in Table C.3 Columns (1) to (9) show the regression coefficients estimated for each day using our preferred specification that includes all conditioning variables.

Table C.2: OLS Estimates - Box Office Decay

| | log(decay) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | **1st Sat.** | **1st Sun.** | **2nd Fri.** | **2nd Sat.** | **2nd Sun** | **3rd Fri.** | **3rd Sat.** | **3rd Sun** |
| log(pre-release volume) | 0.03 | 0.02 | −0.002 | −0.06 | −0.06 | −0.24 | −0.26 | −0.26 |
| | (0.03) | (0.04) | (0.04) | (0.08) | (0.07) | (0.07) | (0.15) | (0.16) |
| log(post-release volume) | −0.08 | −0.05 | −0.03 | 0.01 | 0.01 | 0.31 | 0.32 | 0.32 |
| | (0.03) | (0.04) | (0.04) | (0.08) | (0.07) | (0.07) | (0.14) | (0.15) |
| log(post-release sentiment) | 0.03 | 0.03 | 0.06 | 0.07 | 0.07 | 0.11 | 0.10 | 0.10 |
| | (0.01) | (0.02) | (0.02) | (0.03) | (0.03) | (0.03) | (0.06) | (0.06) |
| log(pre-release adspend) | 0.04 | −0.02 | 0.01 | −0.02 | −0.02 | 0.02 | −0.03 | −0.03 |
| | (0.02) | (0.03) | (0.03) | (0.06) | (0.05) | (0.05) | (0.10) | (0.11) |
| log(post-release adspend) | −0.01 | 0.02 | 0.02 | 0.03 | 0.03 | 0.04 | 0.06 | 0.06 |
| | (0.01) | (0.01) | (0.01) | (0.03) | (0.03) | (0.03) | (0.07) | (0.07) |
| Movie Characteristics | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Month Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| N | 158 | 158 | 158 | 158 | 158 | 157 | 157 | 157 |
| Adj. R-squared | 0.57 | 0.46 | 0.40 | 0.47 | 0.47 | 0.35 | 0.36 | 0.36 |

*Notes:* Parameter estimates from OLS regression of box office decay on Twitter variables, advertising spending and controls. HC1 robust standard errors in parentheses. Box office decay for movie $j$ on day $t$ defined as $\text{decay}_{jt} = \text{box office}_{jt}/\text{box office}_{j0}$, where $t = 0$ is the opening day of release. pre-release volume measures the volume of tweets posted about movie $j$ over the pre-release period and post-release volume measures the amount of tweets from posted from the movie's opening night until day $t-1$. post-release sentiment is the ratio of positive to negative tweets posted from the opening day of release up to and including day $t-1$. pre-release ad spend is the sum of advertising spending in the month leading up to release, whilst post-release ad spend measures advertising spend which has occurred from the opening night until day $t-1$. metascore is again the weighted average of critic reviews and consumer review are fixed effects capturing the CinemaScore of a movie. $X_j$ contains time invariant movie characteristics, including the production budget, whether the movie is part of a series, actor starpower and genre fixed effects. $Z_t$ are month fixed effects which absorb variations in demand due to seasonality across months and an indicator day $t$ is part of a long weekend. Like we saw with opening night box office, pre-release sentiment is also uncorrelated with box office earnings over subsequent periods. Including pre-release sentiment has no influence on the empirical results documented in this section.

Table C.3: OLS Estimates - Box Office Revenue

| | log(box office) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | **1st Fri.** | **1st Sat.** | **1st Sun.** | **2nd Fri.** | **2nd Sat.** | **2nd Sun** | **3rd Fri.** | **3rd Sat.** | **3rd Sun** |
| log(pre-release volume) | 0.39 | −0.12 | −0.21 | −0.28 | −0.35 | −0.35 | −0.52 | −0.55 | −0.55 |
| | (0.05) | (0.09) | (0.09) | (0.09) | (0.13) | (0.12) | (0.12) | (0.21) | (0.22) |
| log(pre-release sentiment) | 0.01 | | | | | | | | |
| | (0.06) | | | | | | | | |
| log(post-release volume) | | 0.47 | 0.58 | 0.66 | 0.69 | 0.69 | 0.99 | 1.00 | 1.00 |
| | | (0.10) | (0.10) | (0.10) | (0.14) | (0.13) | (0.13) | (0.20) | (0.20) |
| log(post-release sentiment) | | 0.01 | 0.04 | 0.07 | 0.08 | 0.08 | 0.09 | 0.08 | 0.08 |
| | | (0.05) | (0.05) | (0.05) | (0.06) | (0.05) | (0.05) | (0.09) | (0.09) |
| log(pre-release adspend) | 0.06 | 0.20 | 0.11 | 0.02 | −0.01 | −0.01 | −0.02 | −0.07 | −0.07 |
| | (0.06) | (0.08) | (0.07) | (0.07) | (0.09) | (0.08) | (0.08) | (0.13) | (0.13) |
| log(post-release adspend) | | −0.08 | −0.03 | 0.04 | 0.05 | 0.05 | 0.11 | 0.12 | 0.12 |
| | | (0.05) | (0.05) | (0.05) | (0.06) | (0.05) | (0.05) | (0.11) | (0.11) |
| log(production budget) | 0.25 | 0.25 | 0.30 | 0.25 | 0.26 | 0.26 | 0.17 | 0.19 | 0.19 |
| | (0.08) | (0.08) | (0.07) | (0.07) | (0.09) | (0.08) | (0.08) | (0.14) | (0.14) |
| Movie Characteristics | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Month Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| N | 158 | 158 | 158 | 158 | 158 | 158 | 157 | 157 | 157 |
| Adj. R-squared | 0.70 | 0.76 | 0.80 | 0.73 | 0.77 | 0.77 | 0.65 | 0.66 | 0.66 |

*Notes:* Parameter estimates from OLS regression of box office revenue (in millions of USD) on Twitter variables, advertising spending and controls. HC1 robust standard errors in parentheses. Variables defined as in Table C.2.

# D Computing Market Shares

The market share of movie $j$ in week $t$ is computed from box office sales data. We combine the box office sales data with information on the average price of movie tickets in the US in year $t$, $p_t$, and US population size estimates, $M_t$.[80] The market share of movie $j$ in time $t$ is then:

$$s_{jt} = \frac{1}{M_t} \frac{\text{box office}_{jt}}{p_t}$$

The market share of the outside good, $s_{0t}$ is then computed as:

$$s_{0t} = 1 - \sum_{j=1}^{J} s_{jt} \mathbb{1}_{jt} \{\text{In Cinema}\}$$

where $\mathbb{1}_{jt} \{\text{In Cinema}\}$ is an indicator variable that takes the value of one when movie $j$ is available in cinemas in period $t$.

[80]Average movie ticket prices are obtained from Box Office Mojo. Our population estimate uses the non-seasonally adjusted resident population obtained from the Bureau of Economic Analysis. The annual estimates for 2014 and 2015 are an average of the BEA's monthly time series for each year and are taken from the Federal Reserve's Economic Data (FRED) database.

# E Supplementary Demand Model Estimates

Table E.1: Nested Logit Demand - OLS Estimates

| | Opening Weekend | Post Opening | Time Invariant |
|---|---|---|---|
| **Time-Varying Parameters** | | | |
| $\text{Volume}_{pre}$ | 0.326 | -0.197 | |
| | (0.139) | (0.177) | |
| $\text{Volume}_{post}$ | | 0.592 | |
| | | (0.254) | |
| $\text{Sentiment}_{pre}$ | -0.002 | -0.029 | |
| | (0.006) | (0.02) | |
| $\text{Sentiment}_{post}$ | | 0.049 | |
| | | (0.023) | |
| $\text{Ad Spend}_{pre}$ | 0.002 | 0.008 | |
| | (0.006) | (0.007) | |
| $\text{Ad Spend}_{post}$ | | 0.074 | |
| | | (0.04) | |
| **Exp. Performance Tier FE** | | | |
| $\theta_{(25,50]}$ | | | 0.91 |
| | | | (0.143) |
| $\theta_{(50,75]}$ | | | 1.308 |
| | | | (0.148) |
| $\theta_{(75,100]}$ | | | 1.652 |
| | | | (0.173) |
| $\theta_{(100,200]}$ | | | 1.835 |
| | | | (0.181) |
| $\theta_{(200,\infty]}$ | | | 2.222 |
| | | | (0.252) |
| **Nest Subst. Parameter** | | | |
| $\rho$ | | | 0.213 |
| | | | (0.065) |
| **Supplementary FE** | | | |
| CinemaScore FE | | | Yes |
| Sequel × Day of Release FE | | | Yes |
| Genre × Day of Release FE | | | Yes |
| Calendar Week FE | | | Yes |

*Notes:* Key parameter estimates from OLS estimation of nested logit demand model with standard errors clustered by movie reported in parentheses. Results should be compared to Table 2 which estimates the same model but instruments for $\rho$ using the number of movies in a genre. Additional parameters for CinemaScore FE, Sequel × Day of Release FE Genre × Day of Release FE, Calendar Week FE, actor starpower, production budget and critic reviews are estimated but not reported.

Table E.2: IV Nested Logit Demand Estimates with Movie Fixed Effects

| | Post Opening | Time Invariant |
|---|---|---|
| **Time Varying Parameters** | | |
| $\text{Volume}_{pre}$ | -0.279 | |
| | (0.136) | |
| $\text{Volume}_{post}$ | 0.42 | |
| | (0.211) | |
| $\text{Sentiment}_{pre}$ | -0.017 | |
| | (0.01) | |
| $\text{Sentiment}_{post}$ | 0.035 | |
| | (0.012) | |
| $\text{Ad Spend}_{pre}$ | 0.003 | |
| | (0.005) | |
| $\text{Ad Spend}_{post}$ | 0.06 | |
| | (0.029) | |
| **Nest Subst. Parameter** | | |
| $\rho$ | | 0.019 |
| | | (0.067) |
| **Supplementary FE** | | |
| Sequel × Day of Release FE | | Yes |
| Genre × Day of Release FE | | Yes |
| Calendar Week FE | | Yes |

*Notes:* Key parameter estimates from IV estimation of nested logit demand model with movie fixed effects. Standard errors clustered by movie reported in parentheses. The inclusion of movie fixed effects means we cannot estimate parameters for pre-release measures because they are time invariant and absorbed into the fixed effect. Results on parameters that can be estimated in both models should be compared to Table 2 which approximates the movie fixed effects with Expected Performance Tier Fixed Effects. As discussed in Section 5.2, we instrument for $\rho$ using the number of movies in a genre. Additional parameters for CinemaScore FE, Sequel × Day of Release FE Genre × Day of Release FE, Calendar Week FE, actor starpower, production budget and critic reviews are estimated but not reported.

¡!–

Table E.3: OLS Nested Logit Demand Estimates with Movie Fixed Effects

| | Post Opening | Time Invariant |
|---|---|---|
| **Time Varying Parameters** | | |
| Volume$_{pre}$ | -0.255 | |
| | (0.137) | |
| Volume$_{post}$ | 0.355 | |
| | (0.212) | |
| Sentiment$_{pre}$ | -0.018 | |
| | (0.01) | |
| Sentiment$_{post}$ | 0.033 | |
| | (0.012) | |
| Ad Spend$_{pre}$ | 0.004 | |
| | (0.005) | |
| Ad Spend$_{post}$ | 0.055 | |
| | (0.028) | |
| **Nest Subst. Parameter** | | |
| $\rho$ | | 0.103 |
| | | (0.055) |
| **Supplementary FE** | | |
| Sequel × Day of Release FE | | Yes |
| Genre × Day of Release FE | | Yes |
| Calendar Week FE | | Yes |

*Notes:* Here are some table notes

−¿

Table E.4: IV Nested Logit Demand Estimates with Time-Varying Expected Performance Tier FE

| | Opening Weekend | Post Opening | Time Invariant |
|---|---|---|---|
| **Time Varying Parameters** | | | |
| $\text{Volume}_{pre}$ | 0.243 | -0.219 | |
| | (0.119) | (0.195) | |
| $\text{Volume}_{post}$ | | 0.697 | |
| | | (0.275) | |
| $\text{Sentiment}_{pre}$ | -0.004 | -0.031 | |
| | (0.007) | (0.02) | |
| $\text{Sentiment}_{post}$ | | 0.055 | |
| | | (0.023) | |
| Ad $\text{Spend}_{pre}$ | 0.006 | 0.006 | |
| | (0.006) | (0.007) | |
| Ad $\text{Spend}_{post}$ | | 0.081 | |
| | | (0.04) | |
| **Segment Fixed Effects** | | | |
| $\theta_{(0,25]}$ | -2.394 | -3.996 | |
| | (0.255) | (0.327) | |
| $\theta_{(25,50]}$ | -1.682 | -2.949 | |
| | (0.238) | (0.306) | |
| $\theta_{(50,75]}$ | -1.227 | -2.515 | |
| | (0.24) | (0.321) | |
| $\theta_{(75,100]}$ | -0.956 | -2.106 | |
| | (0.226) | (0.313) | |
| $\theta_{(100,200]}$ | -0.642 | -2.012 | |
| | (0.187) | (0.282) | |
| $\theta_{(200,\infty]}$ | | -1.742 | |
| | | (0.297) | |
| **Nest Subst. Parameter** | | | |
| $\rho$ | | | 0.111 |
| | | | (0.066) |
| **Supplementary FE** | | | |
| CinemaScore FE | | | Yes |
| Sequel × Day of Release FE | | | Yes |
| Genre × Day of Release FE | | | Yes |
| Calendar Week FE | | | Yes |

*Notes:* Key parameter estimates from IV estimation of nested logit demand model with standard errors clustered by movie reported in parentheses. Expected Performance Tier Fixed Effects are modelled to differ between opening and post-opening weekend. Results should be compared to Table 2 which restricts the Expected Performance Tier Fixed Effects to be constant over the estimation window. As discussed in Section 5.2, we instrument for $\rho$ using the number of movies in a genre. Additional parameters for CinemaScore FE, Sequel × Day of Release FE Genre × Day of Release FE, Calendar Week FE, actor starpower, production budget and critic reviews are estimated but not reported.

Table E.5: IV Nested Logit Demand Estimates with Time-Varying CinemaScore FE

| | Opening Weekend | Post Opening | Time Invariant |
|---|---|---|---|
| **Time Varying Parameters** | | | |
| $\text{Volume}_{pre}$ | 0.349 | -0.208 | |
| | (0.144) | (0.184) | |
| $\text{Volume}_{post}$ | | 0.575 | |
| | | (0.257) | |
| $\text{Sentiment}_{pre}$ | -0.002 | -0.028 | |
| | (0.006) | (0.02) | |
| $\text{Sentiment}_{post}$ | | 0.048 | |
| | | (0.023) | |
| $\text{Ad Spend}_{pre}$ | 0.004 | 0.009 | |
| | (0.006) | (0.007) | |
| $\text{Ad Spend}_{post}$ | | 0.072 | |
| | | (0.04) | |
| **Segment Fixed Effects** | | | |
| $\theta_{(25,50]}$ | | | 0.936 |
| | | | (0.15) |
| $\theta_{(50,75]}$ | | | 1.387 |
| | | | (0.158) |
| $\theta_{(75,100]}$ | | | 1.743 |
| | | | (0.182) |
| $\theta_{(100,200]}$ | | | 1.917 |
| | | | (0.184) |
| $\theta_{(200,\infty]}$ | | | 2.323 |
| | | | (0.253) |
| **Nest Subst. Parameter** | | | |
| $\rho$ | | | 0.102 |
| | | | (0.065) |
| **Supplementary FE** | | | |
| CinemaScore FE | Yes | Yes | No |
| Sequel $\times$ Day of Release FE | | | Yes |
| Genre $\times$ Day of Release FE | | | Yes |
| Calendar Week FE | | | Yes |

*Notes:* Key parameter estimates from IV estimation of nested logit demand model with standard errors clustered by movie reported in parentheses. CinemaScore Fixed Effects are modelled to differ between opening and post-opening weekend. Results should be compared to Table 2 which restricts the CinemaScore Fixed Effects to be constant over the estimation window. As discussed in Section 5.2, we instrument for $\rho$ using the number of movies in a genre. Additional parameters for CinemaScore FE, Sequel $\times$ Day of Release FE Genre $\times$ Day of Release FE, Calendar Week FE, actor starpower, production budget and critic reviews are estimated but not reported.

Table E.6: IV Nested Logit Demand Estimates with Time-Varying Expected Performance Tier and CinemaScore FE

| | Opening Weekend | Post Opening | Time Invariant |
|---|---|---|---|
| **Time Varying Parameters** | | | |
| $\text{Volume}_{pre}$ | 0.28 | -0.226 | |
| | (0.124) | (0.193) | |
| $\text{Volume}_{post}$ | | 0.656 | |
| | | (0.276) | |
| $\text{Sentiment}_{pre}$ | -0.002 | -0.03 | |
| | (0.007) | (0.02) | |
| $\text{Sentiment}_{post}$ | | 0.051 | |
| | | (0.023) | |
| $\text{Ad Spend}_{pre}$ | 0.008 | 0.006 | |
| | (0.006) | (0.007) | |
| $\text{Ad Spend}_{post}$ | | 0.08 | |
| | | (0.04) | |
| **Segment Fixed Effects** | | | |
| $\theta_{(0,25]}$ | -2.436 | -3.865 | |
| | (0.247) | (0.342) | |
| $\theta_{(25,50]}$ | -1.733 | -2.812 | |
| | (0.233) | (0.324) | |
| $\theta_{(50,75]}$ | -1.268 | -2.382 | |
| | (0.233) | (0.34) | |
| $\theta_{(75,100]}$ | -1.003 | -1.973 | |
| | (0.217) | (0.339) | |
| $\theta_{(100,200]}$ | -0.661 | -1.898 | |
| | (0.185) | (0.312) | |
| $\theta_{(200,\infty]}$ | | -1.642 | |
| | | (0.319) | |
| **Nest Subst. Parameter** | | | |
| $\rho$ | | | 0.112 |
| | | | (0.065) |
| **Supplementary FE** | | | |
| CinemaScore FE | Yes | Yes | No |
| Sequel × Day of Release FE | | | Yes |
| Genre × Day of Release FE | | | Yes |
| Calendar Week FE | | | Yes |

*Notes:* Key parameter estimates from IV estimation of nested logit demand model with standard errors clustered by movie reported in parentheses. Expected Performance Tier and CinemaScore Fixed Effects are modelled to differ between opening and post-opening weekend. Results should be compared to Table 2 which restricts the Expected Performance Tier and CinemaScore Fixed Effects to be constant over the estimation window. As discussed in Section 5.2, we instrument for $\rho$ using the number of movies in a genre. Additional parameters for CinemaScore FE, Sequel × Day of Release FE Genre × Day of Release FE, Calendar Week FE, actor starpower, production budget and critic reviews are estimated but not reported.

Table E.7: IV Nested Logit Demand Estimates with Raw CinemaScore Grades as FE

| | Opening Weekend | Post Opening | Time Invariant |
|---|---|---|---|
| **Time Varying Parameters** | | | |
| $\text{Volume}_{pre}$ | 0.363 | -0.196 | |
| | (0.139) | (0.184) | |
| $\text{Volume}_{post}$ | | 0.707 | |
| | | (0.251) | |
| $\text{Sentiment}_{pre}$ | -0.003 | -0.026 | |
| | (0.006) | (0.019) | |
| $\text{Sentiment}_{post}$ | | 0.044 | |
| | | (0.022) | |
| $\text{Ad Spend}_{pre}$ | 0.003 | 0.011 | |
| | (0.006) | (0.007) | |
| $\text{Ad Spend}_{post}$ | | 0.055 | |
| | | (0.036) | |
| **Segment Fixed Effects** | | | |
| $\theta_{(25,50]}$ | | | 0.904 |
| | | | (0.143) |
| $\theta_{(50,75]}$ | | | 1.408 |
| | | | (0.149) |
| $\theta_{(75,100]}$ | | | 1.715 |
| | | | (0.178) |
| $\theta_{(100,200]}$ | | | 1.895 |
| | | | (0.181) |
| $\theta_{(200,\infty]}$ | | | 2.393 |
| | | | (0.263) |
| **Nest Subst. Parameter** | | | |
| $\rho$ | | | 0.114 |
| | | | (0.064) |
| **Supplementary FE** | | | |
| Raw CinemaScore FE | | | Yes |
| Sequel × Day of Release FE | | | Yes |
| Genre × Day of Release FE | | | Yes |
| Calendar Week FE | | | Yes |

*Notes:* Key parameter estimates from IV estimation of nested logit demand model with standard errors clustered by movie reported in parentheses. Raw CinemaScore grades A+, A, A-, B+, B, B-, C+, C, C-, D are included as Fixed Effects. Results should be compared to Table 2 which condenses the CinemaScore Fixed Effects be A, B and "C or lower". As discussed in Section 5.2, we instrument for $\rho$ using the number of movies in a genre. Additional parameters for CinemaScore FE, Sequel × Day of Release FE Genre × Day of Release FE, Calendar Week FE, actor starpower, production budget and critic reviews are estimated but not reported.

Table E.8: IV Nested Logit Demand Estimates with Alternate CinemaScore Aggregation as FE

| | Opening Weekend | Post Opening | Time Invariant |
|---|---|---|---|
| **Time Varying Parameters** | | | |
| $\text{Volume}_{pre}$ | 0.344 | -0.192 | |
| | (0.139) | (0.188) | |
| $\text{Volume}_{post}$ | | 0.599 | |
| | | (0.259) | |
| $\text{Sentiment}_{pre}$ | -0.002 | -0.028 | |
| | (0.006) | (0.02) | |
| $\text{Sentiment}_{post}$ | | 0.049 | |
| | | (0.023) | |
| $\text{Ad Spend}_{pre}$ | 0.004 | 0.01 | |
| | (0.006) | (0.007) | |
| $\text{Ad Spend}_{post}$ | | 0.072 | |
| | | (0.041) | |
| **Segment Fixed Effects** | | | |
| $\theta_{(25,50]}$ | | | 0.946 |
| | | | (0.154) |
| $\theta_{(50,75]}$ | | | 1.393 |
| | | | (0.16) |
| $\theta_{(75,100]}$ | | | 1.748 |
| | | | (0.181) |
| $\theta_{(100,200]}$ | | | 1.937 |
| | | | (0.184) |
| $\theta_{(200,\infty]}$ | | | 2.355 |
| | | | (0.254) |
| **Nest Subst. Parameter** | | | |
| $\rho$ | | | 0.1 |
| | | | (0.066) |
| **Supplementary FE** | | | |
| Alt. CinemaScore FE | | | Yes |
| Sequel × Day of Release FE | | | Yes |
| Genre × Day of Release FE | | | Yes |
| Calendar Week FE | | | Yes |

*Notes:* Key parameter estimates from IV estimation of nested logit demand model with standard errors clustered by movie reported in parentheses. CinemaScore grades are reduced to A+, A, B and "C or lower" to allow for A+ movies to have a different level of offline WoM. Results should be compared to Table 2 which condenses the CinemaScore Fixed Effects be A, B and "C or lower". As discussed in Section 5.2, we instrument for $\rho$ using the number of movies in a genre. Additional parameters for CinemaScore FE, Sequel × Day of Release FE Genre × Day of Release FE, Calendar Week FE, actor starpower, production budget and critic reviews are estimated but not reported.

Table E.9: IV Nested Logit Demand Estimates with HSX Predictions

| | Opening Weekend | Post Opening | Time Invariant |
|---|---|---|---|
| **Time-Varying Parameters** | | | |
| $\text{Volume}_{pre}$ | 0.193 | -0.355 | |
| | (0.114) | (0.183) | |
| $\text{Volume}_{post}$ | | 0.713 | |
| | | (0.262) | |
| $\text{Sentiment}_{pre}$ | -0.003 | -0.029 | |
| | (0.006) | (0.021) | |
| $\text{Sentiment}_{post}$ | | 0.053 | |
| | | (0.023) | |
| $\text{Ad Spend}_{pre}$ | 0.004 | 0.01 | |
| | (0.006) | (0.007) | |
| $\text{Ad Spend}_{post}$ | | 0.078 | |
| | | (0.038) | |
| **Exp. Performance Tier FE** | | | |
| $\theta_{(25,50]}$ | | | 0.899 |
| | | | (0.151) |
| $\theta_{(50,75]}$ | | | 1.298 |
| | | | (0.161) |
| $\theta_{(75,100]}$ | | | 1.601 |
| | | | (0.185) |
| $\theta_{(100,200]}$ | | | 1.702 |
| | | | (0.193) |
| $\theta_{(200,\infty]}$ | | | 1.738 |
| | | | (0.296) |
| **Nest Subst. Parameter** | | | |
| $\rho$ | | | 0.098 |
| | | | (0.065) |
| **Additional Parameters** | | | |
| HSX Pred. Performance | | | 0.003 |
| | | | (0.001) |
| **Supplementary FE** | | | |
| CinemaScore FE | | | Yes |
| Sequel × Day of Release FE | | | Yes |
| Genre × Day of Release FE | | | Yes |
| Calendar Week FE | | | Yes |

*Notes:* Key parameter estimates from IV estimation of nested logit demand model with standard errors clustered by movie reported in parentheses. We include the average HSX closing price over the week pre-release as an additional control. Results should be compared to Table 2. As discussed in Section 5.2, we instrument for $\rho$ using the number of movies in a genre. Additional parameters for CinemaScore FE, Sequel × Day of Release FE Genre × Day of Release FE, Calendar Week FE, actor starpower, production budget and critic reviews are estimated but not reported.

Table E.10: IV Nested Logit Demand Estimates with HSX Updates

| | Opening Weekend | Post Opening | Time Invariant |
|---|---|---|---|
| **Time-Varying Parameters** | | | |
| $\text{Volume}_{pre}$ | 0.184 | 0.408 | |
| | (0.15) | (0.171) | |
| $\text{Volume}_{post}$ | | -0.473 | |
| | | (0.22) | |
| $\text{Sentiment}_{pre}$ | -0.005 | -0.031 | |
| | (0.006) | (0.02) | |
| $\text{Sentiment}_{post}$ | | 0.054 | |
| | | (0.022) | |
| $\text{Ad Spend}_{pre}$ | -0.003 | 0.009 | |
| | (0.005) | (0.007) | |
| $\text{Ad Spend}_{post}$ | | 0.053 | |
| | | (0.037) | |
| **Exp. Performance Tier FE** | | | |
| $\theta_{(25,50]}$ | | | 0.986 |
| | | | (0.147) |
| $\theta_{(50,75]}$ | | | 1.496 |
| | | | (0.157) |
| $\theta_{(75,100]}$ | | | 1.897 |
| | | | (0.171) |
| $\theta_{(100,200]}$ | | | 2.1 |
| | | | (0.19) |
| $\theta_{(200,\infty]}$ | | | 2.511 |
| | | | (0.251) |
| **Nest Subst. Parameter** | | | |
| $\rho$ | | | 0.1 |
| | | | (0.065) |
| **Additional Parameters** | | | |
| $\Delta$ HSX Price | | | 0.006 |
| | | | (0.001) |
| **Supplementary FE** | | | |
| CinemaScore FE | | | Yes |
| Sequel × Day of Release FE | | | Yes |
| Genre × Day of Release FE | | | Yes |
| Calendar Week FE | | | Yes |

*Notes:* Key parameter estimates from IV estimation of nested logit demand model with standard errors clustered by movie reported in parentheses. We include the change in the HSX price, defined as the difference between the price on the close of a Thursdays trading for a given week from its pre-release prediction as an additional control. Results should be compared to Table 2. As discussed in Section 5.2, we instrument for $\rho$ using the number of movies in a genre. Additional parameters for CinemaScore FE, Sequel × Day of Release FE Genre × Day of Release FE, Calendar Week FE, actor starpower, production budget and critic reviews are estimated but not reported.

Table E.11: IV Nested Logit Demand Estimates with HSX Updates Interacted with Sign of Change

| | Opening Weekend | Post Opening | Time Invariant |
|---|---|---|---|
| **Time-Varying Parameters** | | | |
| $\text{Volume}_{pre}$ | 0.202 | 0.435 | |
| | (0.159) | (0.187) | |
| $\text{Volume}_{post}$ | | -0.478 | |
| | | (0.217) | |
| $\text{Sentiment}_{pre}$ | -0.005 | -0.031 | |
| | (0.006) | (0.02) | |
| $\text{Sentiment}_{post}$ | | 0.054 | |
| | | (0.022) | |
| $\text{Ad Spend}_{pre}$ | -0.003 | 0.009 | |
| | (0.005) | (0.007) | |
| $\text{Ad Spend}_{post}$ | | 0.051 | |
| | | (0.037) | |
| **Exp. Performance Tier FE** | | | |
| $\theta_{(25,50]}$ | | | 0.995 |
| | | | (0.148) |
| $\theta_{(50,75]}$ | | | 1.516 |
| | | | (0.16) |
| $\theta_{(75,100]}$ | | | 1.907 |
| | | | (0.172) |
| $\theta_{(100,200]}$ | | | 2.116 |
| | | | (0.196) |
| $\theta_{(200,\infty]}$ | | | 2.574 |
| | | | (0.256) |
| **Nest Subst. Parameter** | | | |
| $\rho$ | | | 0.097 |
| | | | (0.065) |
| **Additional Parameters** | | | |
| Negative $\Delta$ HSX Price | | | 0.008 |
| | | | (0.003) |
| Positive $\Delta$ HSX Price | | | 0.005 |
| | | | (0.001) |
| **Supplementary FE** | | | |
| CinemaScore FE | | | Yes |
| Sequel $\times$ Day of Release FE | | | Yes |
| Genre $\times$ Day of Release FE | | | Yes |
| Calendar Week FE | | | Yes |

*Notes:* Key parameter estimates from IV estimation of nested logit demand model with standard errors clustered by movie reported in parentheses. We include the change in the HSX price, defined as the difference between the price on the close of a Thursdays trading for a given week from its pre-release prediction as an additional control. These changes are interacted with the sign of the change, so positive and negative changes get their own coefficient. Results should be compared to Table 2. As discussed in Section 5.2, we instrument for $\rho$ using the number of movies in a genre. Additional parameters for CinemaScore FE, Sequel $\times$ Day of Release FE Genre $\times$ Day of Release FE, Calendar Week FE, actor starpower, production budget and critic reviews are estimated but not reported.

# F Demand Estimates with Heterogeneity

Table F.1: Nested Logit Demand Estimates with Franchise Heterogeneity

| | Opening Weekend | Post Opening |
|---|---|---|
| **Franchise = True** | | |
| $\text{Volume}_{pre}$ | 0.512 | 0.224 |
| | (0.186) | (0.228) |
| $\text{Volume}_{post}$ | | 0.184 |
| | | (0.223) |
| $\text{Sentiment}_{pre}$ | 0.009 | 0.016 |
| | (0.006) | (0.005) |
| $\text{Sentiment}_{post}$ | | 0 |
| | | (0.021) |
| $\text{Ad Spend}_{pre}$ | 0.012 | 0.016 |
| | (0.012) | (0.012) |
| $\text{Ad Spend}_{post}$ | | -0.01 |
| | | (0.028) |
| **Franchise = False** | | |
| $\text{Volume}_{pre}$ | 0.175 | -1.194 |
| | (0.118) | (0.324) |
| $\text{Volume}_{post}$ | | 2.503 |
| | | (0.658) |
| $\text{Sentiment}_{pre}$ | -0.012 | -0.086 |
| | (0.01) | (0.013) |
| $\text{Sentiment}_{post}$ | | 0.129 |
| | | (0.019) |
| $\text{Ad Spend}_{pre}$ | 0.004 | 0.008 |
| | (0.006) | (0.007) |
| $\text{Ad Spend}_{post}$ | | 0.167 |
| | | (0.046) |

*Notes:* Key parameter estimates from IV estimation of nested logit demand model with franchise heterogeneity. Standard errors clustered by movie reported in parentheses. Parameters for tweet volume, tweet sentiment and advertising expenditure are modelled to differ between franchise and non-franchise movies. Franchise movies are defined as movies that are a sequel of another, with an interdependant story line, or where characters of a movie are based on an existing film. Additional parameters for Expected Performance Tier FE, the nested logit substitution parameter, $\rho$, CinemaScore FE, Sequel $\times$ Day of Release FE Genre $\times$ Day of Release FE, Calendar Week FE, actor starpower, production budget and critic reviews are estimated but not reported. Own demand elasticities that stem from these parameter estimates are in Table 4.

Table F.2: Nested Logit Demand Estimates with Expected Performance Tier Heterogeneity

| | Opening Weekend | Post Opening |
|---|---|---|
| **Exp. Performance Tier = Large** | | |
| $\text{Volume}_{pre}$ | 0.412 | 0.229 |
| | (0.206) | (0.232) |
| $\text{Volume}_{post}$ | | 0.182 |
| | | (0.277) |
| $\text{Sentiment}_{pre}$ | 0.022 | 0.237 |
| | (0.036) | (0.106) |
| $\text{Sentiment}_{post}$ | | -0.204 |
| | | (0.085) |
| $\text{Ad Spend}_{pre}$ | 0.025 | 0.043 |
| | (0.022) | (0.025) |
| $\text{Ad Spend}_{post}$ | | -0.039 |
| | | (0.026) |
| **Exp. Performance Tier = Medium** | | |
| $\text{Volume}_{pre}$ | 0.028 | -0.546 |
| | (0.105) | (0.29) |
| $\text{Volume}_{post}$ | | 0.965 |
| | | (0.495) |
| $\text{Sentiment}_{pre}$ | -0.001 | 0.019 |
| | (0.006) | (0.007) |
| $\text{Sentiment}_{post}$ | | -0.038 |
| | | (0.019) |
| $\text{Ad Spend}_{pre}$ | -0.014 | 0.009 |
| | (0.008) | (0.009) |
| $\text{Ad Spend}_{post}$ | | 0.055 |
| | | (0.061) |
| **Exp. Performance Tier = Small** | | |
| $\text{Volume}_{pre}$ | 1.159 | 1.085 |
| | (0.405) | (0.683) |
| $\text{Volume}_{post}$ | | 1.203 |
| | | (1.071) |
| $\text{Sentiment}_{pre}$ | -0.011 | -0.088 |
| | (0.01) | (0.013) |
| $\text{Sentiment}_{post}$ | | 0.127 |
| | | (0.018) |
| $\text{Ad Spend}_{pre}$ | 0.01 | 0.014 |
| | (0.007) | (0.009) |
| $\text{Ad Spend}_{post}$ | | 0.114 |
| | | (0.07) |

*Notes:* Key parameter estimates from IV estimation of nested logit demand model with Expected Performance Tier heterogeneity. Standard errors clustered by movie reported in parentheses. Parameters for tweet volume, tweet sentiment and advertising expenditure are modelled to differ between Expected Performance Tiers. See footnote 67 for definitions of performance tiers. Additional parameters for Expected Performance Tier FE, the nested logit substitution parameter, $\rho$, CinemaScore FE, Sequel × Day of Release FE Genre × Day of Release FE, Calendar Week FE, actor starpower, production budget and critic reviews are estimated but not reported. Own

# References


Ainslie, A., Drèze, X. & Zufryden, F. (2005), 'Modeling movie life cycles and market share', *Marketing Science* **24**(3), 508–517.

American Marketing Association (2013), *Social Media Success Metrics*.
**URL:** *https://www.ama.org/ECDFileRepository/Special-Report-Social-Media-Success-Metrics.pdf*

Apala, K. R., Jose, M., Motnam, S., Chan, C.-C., Liszka, K. J. & de Gregorio, F. (2013), Prediction of movies box office performance using social media, *in* 'Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining', ACM, pp. 1209–1214.

Babić Rosario, A., Sotgiu, F., De Valck, K. & Bijmolt, T. H. (2016), 'The effect of electronic word of mouth on sales: A meta-analytic review of platform, product, and metric factors', *Journal of Marketing Research* **53**(3), 297–318.

Baek, H., Ahn, J. & Oh, S. (2014), 'Impact of tweets on box office revenue: focusing on when tweets are written', *ETRI Journal* **36**(4), 581–590.

Baek, H., Oh, S., Yang, H.-D. & Ahn, J. (2017), 'Electronic word-of-mouth, box office revenue and social media', *Electronic Commerce Research and Applications* **22**, 13–23.

Baer, N., Messina, A. & Murri, R. (2018), 'Elasticluster'.
**URL:** *http://elasticluster.readthedocs.io/en/latest/index.html*

Berry, S. T. (1994), 'Estimating discrete-choice models of product differentiation', *The RAND Journal of Economics* pp. 242–262.

Bollen, J., Mao, H. & Zeng, X. (2011), 'Twitter mood predicts the stock market', *Journal of computational science* **2**(1), 1–8.

Campbell, A., Mayzlin, D. & Shin, J. (2017), 'Managing buzz', *The RAND Journal of Economics* **48**(1), 203–229.

Campbell, M. C. & Keller, K. L. (2003), 'Brand familiarity and advertising repetition effects', *Journal of consumer research* **30**(2), 292–304.

Cardell, N. S. (1997), 'Variance components structures for the extreme-value and logistic distributions with application to models of heterogeneity', *Econometric Theory* **13**(2), 185–213.

Chevalier, J. A. & Mayzlin, D. (2006), 'The effect of word of mouth on sales: Online book reviews', *Journal of marketing research* **43**(3), 345–354.

Chintagunta, P. K., Gopinath, S. & Venkataraman, S. (2010), 'The effects of online user reviews on movie box office performance: Accounting for sequential rollout and aggregation across local markets', *Marketing Science* **29**(5), 944–957.

Chintagunta, P. K., Kadiyali, V. & Vilcassim, N. J. (2006), 'Endogeneity and simultaneity in competitive pricing and advertising: A logit demand analysis', *The Journal of Business* **79**(6), 2761–2787.

Chu, S.-C. & Kim, Y. (2011), 'Determinants of consumer engagement in electronic word-of-mouth (ewom) in social networking sites', *International journal of Advertising* **30**(1), 47–75.

Dhar, V. & Chang, E. A. (2009), 'Does chatter matter? the impact of user-generated content on music sales', *Journal of Interactive Marketing* **23**(4), 300–307.

Ding, C., Cheng, H. K., Duan, Y. & Jin, Y. (2017), 'The power of the like button: The impact of social media on box office', *Decision Support Systems* **94**, 77–84.

Duan, W., Gu, B. & Whinston, A. B. (2008), 'Do online reviews matter?an empirical investigation of panel data', *Decision support systems* **45**(4), 1007–1016.

Dubé, J.-P., Hitsch, G. J. & Manchanda, P. (2005), 'An empirical model of advertising dynamics', *Quantitative marketing and economics* **3**(2), 107–144.

Einav, L. (2007), 'Seasonality in the us motion picture industry', *The Rand journal of economics* **38**(1), 127–145.

Foutz, N. Z. & Jank, W. (2010), 'Research noteprerelease demand forecasting for motion pictures using functional shape analysis of virtual stock markets', *Marketing Science* **29**(3), 568–579.

Gelper, S., Peres, R. & Eliashberg, J. (2018), 'Talk bursts: The role of spikes in prerelease word-of-mouth dynamics', *Journal of Marketing Research* **55**(6), 801–817.

Gilbert, E. & Hutto, C. (2014), Vader: A parsimonious rule-based model for sentiment analysis of social media text, *in* 'Eighth International Conference on Weblogs and Social Media (ICWSM-14)'.

Gilchrist, D. S. & Sands, E. G. (2016), 'Something to talk about: Social spillovers in movie consumption', *Journal of Political Economy* **124**(5), 1339–1382.

Godes, D. & Mayzlin, D. (2004), 'Using online conversations to study word-of-mouth communication', *Marketing science* **23**(4), 545–560.

Goldsmith, R. E. & Horowitz, D. (2006), 'Measuring motivations for online opinion seeking', *Journal of interactive advertising* **6**(2), 2–14.

Gong, S., Zhang, J., Zhao, P. & Jiang, X. (2018), 'Tweeting as a marketing tool: A field experiment in the tv industry', *Journal of Marketing Research* .

Gopinath, S., Chintagunta, P. K. & Venkataraman, S. (2013), 'Blogs, advertising, and local-market movie box office performance', *Management Science* **59**(12), 2635–2654.

Gordon, B. R. & Hartmann, W. R. (2013), 'Advertising effects in presidential elections', *Marketing Science* **32**(1), 19–35.

Hennig-Thurau, T., Gwinner, K. P., Walsh, G. & Gremler, D. D. (2004), 'Electronic word-of-mouth via consumer-opinion platforms: what motivates consumers to articulate themselves on the internet?', *Journal of interactive marketing* **18**(1), 38–52.

Hennig-Thurau, T., Wiertz, C. & Feldhaus, F. (2015), 'Does twitter matter? the impact of microblogging word of mouth on consumers adoption of new movies', *Journal of the Academy of Marketing Science* **43**(3), 375–394.

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.

Hu, G., Bhargava, P., Fuhrmann, S., Ellinger, S. & Spasojevic, N. (2017), 'Analyzing users' sentiment towards popular consumer industries and brands on twitter', *arXiv preprint arXiv:1709.07434* .

Ishihara, M. & Moorthy, S. (2018), 'On the dynamics of brand extensions: The case of movies'.

Jain, V. (2013), 'Prediction of movie success using sentiment analysis of tweets', *The International Journal of Soft Computing and Software Engineering* **3**(3), 308–313.

Jansen, B. J., Zhang, M., Sobel, K. & Chowdury, A. (2009), 'Twitter power: Tweets as electronic word of mouth', *Journal of the American Society for Information Science and Technology* **60**(11), 2169–2188.
**URL:** *https://onlinelibrary.wiley.com/doi/abs/10.1002/asi.21149*

Jarvey, N. (2014), 'Twitter survey explains why movie tweets matter', Hollywood Reporter.

Kim, H. & Hanssens, D. M. (2017), 'Advertising and word-of-mouth effects on pre-launch consumer interest and initial sales of experience products', *Journal of Interactive Marketing* **37**, 57–74.

Köster, J. & Rahmann, S. (2012), 'Snakemakea scalable bioinformatics workflow engine', *Bioinformatics* **28**(19), 2520–2522.

Kuehn, J. & Lampe, R. (2018), 'Social learning or network externalities?', *Working Paper* .

Li, H. (2018), 'Alluxio: A virtual distributed file system'.

Li, H., Ghodsi, A., Zaharia, M., Shenker, S. & Stoica, I. (2014), Tachyon: Reliable, memory speed storage for cluster computing frameworks, *in* 'Proceedings of the ACM Symposium on Cloud Computing', ACM, pp. 1–15.

Liu, X., Lee, D. & Srinivasan, K. (2018), Large scale cross-category analysis of consumer review content on sales conversion leveraging deep learning, *in* 'Workshops at the Thirty-Second AAAI Conference on Artificial Intelligence'.

Liu, Y. (2006), 'Word of mouth for movies: Its dynamics and impact on box office revenue', *Journal of marketing* **70**(3), 74–89.

Lovett, M. J. & Staelin, R. (2016), 'The role of paid, earned, and owned media in building entertainment brands: Reminding, informing, and enhancing enjoyment', *Marketing Science* **35**(1), 142–157.

McFadden, D. (1978), 'Modeling the choice of residential location', *Transportation Research Record* (673).

MPAA (2016), Theatrical Market Statistics - 2016, Technical report.

MPAA (2017), Theme Report - A comprehensive analysis and survey of the theatrical and home entertainment environment, Technical report.

Mummalaneni, S., Wang, Y., K., C. P. & Dhar, S. K. (2019), 'Product placement effects on store sales: Evidence from consumer packaged goods'.

Murphy, D. (2014), 44 percent of Twitter users have never tweeted - PC Magazine, Technical report.

Nguyen, L. T., Wu, P., Chan, W., Peng, W. & Zhang, Y. (2012), Predicting collective sentiment dynamics from time-series social media, *in* 'Proceedings of the first international workshop on issues of sentiment discovery and opinion mining', ACM, p. 6.

Nielsen (2013), *Global Trust in Advertising and Brand Messages.*
**URL:** *http://www.eaca.eu/wp-content/uploads/2016/06/Global-Trust-in-Advertising.pdf*

O'Connor, B., Balasubramanyan, R., Routledge, B. R., Smith, N. A. et al. (2010), 'From tweets to polls: Linking text sentiment to public opinion time series.', *Icwsm* **11**(122-129), 1–2.

OpenStack (2018), 'Openstack swift object storage service'.
**URL:** *https://docs.openstack.org/swift/latest/*

Rennhoff, A. D. & Wilbur, K. C. (2011), 'The effectiveness of post-release movie advertising', *International Journal of Advertising* **30**(2), 305–328.

Ribeiro, F. N., Araújo, M., Gonçalves, P., Gonçalves, M. A. & Benevenuto, F. (2016), 'Sentibench-a benchmark comparison of state-of-the-practice sentiment analysis methods', *EPJ Data Science* **5**(1), 1–29.

Robehmed, N. (2015), 'For movies, twitter users are a key audience', Forbes Magazine.

Rossi, P. E. (2014), 'Even the rich can make themselves poor: A critical examination of iv methods in marketing applications', *Marketing Science* **33**(5), 655–672.

Rui, H., Liu, Y. & Whinston, A. (2013), 'Whose and what chatter matters? the effect of tweets on movie sales', *Decision Support Systems* **55**(4), 863–870.

Schweidel, D. A. & Moe, W. W. (2014), 'Listening in on social media: A joint model of sentiment and venue format choice', *Journal of Marketing Research* **51**(4), 387–402.

Seiler, S., Yao, S. & Wang, W. (2017), 'Does online word of mouth increase demand?(and how?) evidence from a natural experiment', *Marketing Science* **36**(6), 838–861.

Shapiro, B. T. (2018*a*), 'Advertising in health insurance markets', *Marketing Science* .

Shapiro, B. T. (2018*b*), 'Positive spillovers and free riding in advertising of prescription pharmaceuticals: The case of antidepressants', *Journal of Political Economy* **126**(1), 381–437.

Si, J., Mukherjee, A., Liu, B., Li, Q., Li, H. & Deng, X. (2013), Exploiting topic based twitter sentiment for stock prediction, *in* 'Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)', Vol. 2, pp. 24–29.

Siegler, M. (2011), At 1.6 Billion Queries Per Day, Twitter Finally Aims To Make Search Personally Relevant, Technical report.

Sinkinson, M. & Starc, A. (2018), 'Ask your doctor? direct-to-consumer advertising of pharmaceuticals', *The Review of Economic Studies* **86**(2), 836–881.

Smith, A. N., Fischer, E. & Yongjian, C. (2012), 'How does brand-related user-generated content differ across youtube, facebook, and twitter?', *Journal of Interactive Marketing* **26**(2), 102 – 113.
**URL:** *http://www.sciencedirect.com/science/article/pii/S1094996812000059*

Song, R., Jang, S. & Cai, G. G. (2016), 'Does advertising indicate product quality? evidence from prelaunch and postlaunch advertising in the movie industry', *Marketing Letters* **27**(4), 791–804.

Song, T., Huang, J., Tan, Y. & Yu, Y. (2019), 'Using user-and marketer-generated content for box office revenue prediction: Differences between microblogging and third-party platforms', *Information Systems Research* .

Spann, M. & Skiera, B. (2003), 'Internet-based virtual stock markets for business forecasting', *Management science* **49**(10), 1310–1326.

Sun, T., Youn, S., Wu, G. & Kuntaraporn, M. (2006), 'Online word-of-mouth (or mouse): An exploration of its antecedents and consequences', *Journal of Computer-Mediated Communication* **11**(4), 1104–1127.

Suslak, M. (2014), 'If it's happening at the movies, it's happening on twitter', The Official Twitter Blog.

Thomas, M. (2017), 'Mass advertising as a natural experiment'.

Toubia, O. & Stephen, A. T. (2013), 'Intrinsic vs. image-related utility in social media: Why do people contribute content to twitter?', *Marketing Science* **32**(3), 368–392.

Tuchman, A. (2018), 'Advertising and demand for addictive goods: The effects of e-cigarette advertising', *Available at SSRN 3182730* .

Ward, T. (2018), 'How much social media is too much', Forbes Magazine.

WOMMA (2013), ‘Marketers Say ”Word of Mouth Marketing” Is More Effective than Traditional Marketing; Forecast Big Increase in Social Media Spending’, PR Newswire.

Wong, F. M. F., Sen, S. & Chiang, M. (2012), Why watching movie tweets won’t tell the whole story?, *in* ‘Proceedings of the 2012 ACM workshop on Workshop on online social networks’, ACM, pp. 61–66.

You, Y., Vadakkepatt, G. G. & Joshi, A. M. (2015), ‘A meta-analysis of electronic word-of-mouth elasticity’, *Journal of Marketing* **79**(2), 19–39.

Zaharia, M., Chowdhury, M., Franklin, M. J., Shenker, S. & Stoica, I. (2010), ‘Spark: Cluster computing with working sets.’, *HotCloud* **10**(10-10), 95.

Zaharia, M., Xin, R. S., Wendell, P., Das, T., Armbrust, M., Dave, A., Meng, X., Rosen, J., Venkataraman, S., Franklin, M. J. et al. (2016), ‘Apache spark: a unified engine for big data processing’, *Communications of the ACM* **59**(11), 56–65.

Zielske, H. A. (1959), ‘The remembering and forgetting of advertising’, *the Journal of Marketing* pp. 239–243.