INSEAD The Business School for the World®

Working Paper
2019/25/EFE

David vs. Goliath in the Digital Age: The Effect of Network Structure and Content on the Adoption of Cultural Products

Jamie Seoyeon Song
INSEAD, seoyeon.song@insead.edu

Martin Gargiulo
INSEAD, martin.gargiulo@insead.edu

Social media platforms have transformed the way with which communication among industry players and relevant audiences can influence adoption of a cultural product. Using Twitter conversations regarding eBooks sold on Amazon.com between 2014 and 2015, this paper studies how producers in the U.S. publishing industry benefit differentially from Twitter network structure and contents. We find that generalists such as the Big Five publishers gain advantage from controversial contents that appeal to the general public whereas peripheral specialists such as indie publishers benefit from network redundancy and consensus of opinion within niche communities. By finding how producers targeting different audiences benefit from online network structure and content, we explore different strategies with which respective players can leverage social media platforms and accrue competitive advantage.

Keywords: Social Network; Online Platform; Diffusion; Twitter; Content Analysis

Introduction

Digital platforms have transformed how organizations interact with their different stakeholders. They have enabled supply-side innovation by facilitating the coordination of downstream complementarities and reducing fixed costs related to setting up individual organizations (Gawer, 2014). They have also enabled the demand-side innovation by lowering marketing costs for sellers and search costs for buyers (Armstrong, 2006; Evans, Hagiu, & Schmalensee, 2006; Rochet & Tirole, 2008; Stiglitz, 1989). In both cases, platforms such as social networking sites with open interface have significantly changed the cost structure of industries, which in turn can affect the power distribution and the competitive relationships among the existing organizations (Benner & Waldfogel, 2016; Greve & Song, 2017).

Social media platforms are a special case in this digital ecosystem, because they provide new and easily accessible ways through which industry players can influence potential consumers. The special status of social media is grounded on several converging reasons. First, social media platforms create new ways in which producers can reach and interact with actual or potential consumers, eliminating barriers that hitherto prevented smaller producers to reach potential customers. Second, because social media platforms also enable interactions *between* customers, they can augment the scope of traditional word-of-mouth marketing, with important consequences for the process through which customers make decisions regarding products. By facilitating interactions between producers and consumers, social media platforms open new ways through which producers can create awareness and acceptance for their products. At the same time, the expansion of customer-to-customer interaction may render the producers more dependent on processes driven by exchanges of information and opinions among consumers, which are largely beyond the control of those producers. The effect of these exchanges is particularly salient for cultural products, because the legitimacy and desirability of these products may depend on intersubjective judgements by the relevant audiences.

The idea that interpersonal influence affects the way people make choices has a long tradition in the social sciences, spanning from early research on voting behavior (Katz & Lazarsfeld, 1955) to recent studies on the effect of interpersonal influence on the decision to download unknown songs (Salganik, Dodds, & Watts, 2006). Scholars contributing to this tradition have shown that individual decisions can be modelled as a function of the opinions or behaviors of relevant others, such as people the focal individual is connected to in the social network or those who occupy a similar position in the structure of that network (Burt, 1987; see Friedkin & Johnsen 2013 for a comprehensive treatment). An important implication of this research is that the characteristics of the social network linking actual or potential customers, as well as the content of the communication that circulates in that network, can have important consequences for the behavior of the customers and thus for the performance of the organizations that depend on that behavior. In particular, scholars have shown that sparse networks increase the probability that a given message would reach a large number of people in the network (e.g., Granovetter, 1973, 1978), whereas the density of connections in the network increases the probability that this same individual would receive several "redundant" messages on a specific topic (Centola & Macy, 2007; Coleman et al., 1966)

Despite this rich research tradition, scholars have not sufficiently explored how the findings from the study of traditional interpersonal network apply to the online world. While some studies suggest that the behavior of people in online networks may be essentially similar to the behavior identified in traditional influence studies (e.g., Salganik et al., 2006), the dramatic increase in the ability to reach a large number of people with negligible marginal costs may substantially alter the trade-off between reach and redundancy of communication. This could have important consequences for the online strategies of producers. For example, should producers take advantage of the negligible marginal cost of targeting additional potential consumers in social media and use them to maximize the number of people they reach, or should they ensure a sufficient level of redundancy to secure these people's attention? A similar tension can affect the content of the messages circulating through social media. Do producers benefit from consistently positive messages about their products in social media, or are they better off by the raised attention generated by controversy around their products? The tension between consistency and attention may be particularly consequential for cultural products, given the importance of intersubjective judgements in the adoption of such products. Do cultural products benefit more from consistently positive evaluations that from the buzz created by controversy in those evaluations? Are these effects similar for all cultural products, or are they contingent on the type of markets targeted by those products?

These are the questions we seek to answer in this paper. We argue that the effects of the structure of the online communication networks and the content of the messages that circulates through those networks on the adoption of cultural products are contingent on the characteristics of the targeted markets, which correlates with the identity of producers. Specifically, we argue that specialist cultural producers targeting "niche" markets whose participants are likely to share common values and interests will benefit from the redundancy of online communication enabled by more densely connected networks. Redundancy, however, will not benefit as much—and it may even not benefit at all—the performance of generalists targeting the mass market. Likewise, we argue that while generalist cultural producers targeting the mass market will benefit from the raised attention caused by controversy in social media around their products, controversy should hurt the performance of specialist cultural producers targeting niche markets. More generally, the more the consumption of a product can be viewed as a symbolic affirmation of belonging to a community of like-minded people, the stronger the negative effect of controversy on product sales.

We test our ideas by studying the effect of the structure and the content of online communication in a digital platform (Twitter) on the sales of electronically published books (eBooks) in the US between 2014 and 2015. The US book publishing industry has traditionally been dominated by the generalists "Big Five" publishers (Hachette, HarperCollins, McMillan, Random House, and Simon & Schuster), whose products aim to a large and differentiated market, but it also contains a large number of specialist "indie" publishers targeting niche markets of people who often espouse values or lifestyles that sets them apart from the mainstream. This industry structure creates an ideal context to test the different effect of network structure and content on the fate of new eBooks, measured in terms of sales through the Amazon website.

Our findings suggest that specialist publishers can gain competitive advantage by leveraging the mutual reassurance that results from redundant messages around their products while seeking to minimize controversy around those products. Communication redundancy, however, does not benefit generalist producers. Because these producers target a large, undifferentiated mainstream market, they are less likely to benefit from the reassurances associated with redundant messages around their products. By the same reason, generalists benefit from the increasing buzz generated by controversy in the online communication about their products, whereas the uncertainty on the appropriateness of their products for their targeted niche generated by controversy hurts specialists. These findings suggest that different type of producers can gain competitive advantage by leveraging social media platforms in different ways, depending on the characteristics of their targeted markets.

Online Communication Networks and the Adoption of Cultural Products

Social media platforms are a particularly effective mean through which producers can reach actual and potential consumers and make them aware of their products. In addition to allowing messages targeted to specific individuals, social media allow for messages to reach third parties that were not intially targeted by the initiators of the communication (Brown, Broderick, & Lee, 2007). This non-exclusive and public nature of the messages implies that media attention can be easily initiated and spilled over to others, sometimes creating a cascade within online communities (Kwak, Lee, Park, & Moon, 2010; Toubiana & Zietsma, 2016). These characteristics make social media an effective way to build awareness and acceptance for new products, bringing the power of traditional "word of mouth" marketing to an entirely different level (Berger, 2014; Oliver & Swan, 1989).

Although social media facilitate interactions between producers and potential consumers, they also introduce new challenges for producers. By drastically lowering the barriers to communication between their members, social media also enable interactions *between* consumers at a scale and scope that vastly exceed the limits of traditional face-to-face communication. While producers may seek to influence the content of the exchanges among members of the relevant audiences, the extent of this influence is necessarily limited by the very nature of online social media, which enable unrestricted exchanges among platform members. Insofar as these exchanges can be conduits for interpersonal influence, they can amplify the impact of the claims made by producers and reinforce the positive evaluation of the product, but they may also undermine those claims by casting doubts about the value of the product.

The study of the effects of interpersonal influence on individual attitudes and behaviors has a long tradition in the social sciences, going from the pioneering work on voting behavior (Katz & Lazarsfeld, 1955) and diffusion of innovations (Coleman, Katz, & Menzel, 1957) to the structural theory proposed by Friedkin and his collaborators to account for influence processes in small groups (Friedkin, 1993; Friedkin & Johnsen, 2013). This research tradition has highlighted the importance of the structure of the social network as well as the content of the messages circulating through this network in shaping the pattern and the speed of the adoption of new ideas and products (see Muller & Peres, 2019 for a recent review). In particular, scholars have shown that interpersonal communication can have a substantial impact on the way people value a cultural product, even *after* they have consumed and evaluated that same product individually (Childress & Friedkin, 2012).

Recent research showing similarities in the structure of online and offline communication networks (Gilbert & Karahalios, 2009; Leskovec & Horvitz, 2008; Rainie & Wellman, 2012) suggests that the influence processes and mechanisms documented in off-line networks may also apply to online networks. From a researcher's viewpoint, there are two important differences between offline and online networks, however. First, online exchanges typically leave "electronic breadcrumbs" (Golder & Macy, 2014) that can enable researchers to obtain data on the timing and the content of those exchanges. This makes it possible to study the effects of network structure and network content on behavior in a scale and scope that was rarely possible in traditional studies. Second, because they can accelerate the spread of information, online networks are more susceptible to "cascade" effects that amplify the effects of influence processes. Such dramatic increase in people's ability to reach others has substantially altered the trade-off between communication reach and redundancy, highlighting the importance of reevaluating previous claims that emphasized the importance of network reach over redundancy (Granovetter, 1973; Onnela et al., 2007; Watts, 1999).

The Effect of Network Structure

The effects of the structure of a social network in facilitating or hindering the spread of ideas, norms, and behaviors has been highlighted in a number of studies (see Centola, 2010; Newman, Barabási, & Watts, 2006; Watts & Strogatz, 1998). Building on Granovetter's (1973) pathbreaking insight that "weak ties" can connect otherwise separated components in a network, most of this research has highlighted the role of sparse networks in facilitating rapid diffusion of knowledge and innovation (e.g., Abrahamson & Rosenkopf, 1997; Burt, 1999; Hedström, Sandell, & Stern, 2000). Modeled after the diffusion of contagious diseases, these studies often assume that it is enough to be sufficiently exposed to a single "infected" source to contract the “virus”. If all what you need to know to buy a particular product is reassurance from *one* of your trusted contacts, redundant ties (that is, knowing about more contacts who also experienced the product) are irrelevant. If anything, redundant ties may reduce adoption, as many of the messages circulating through the network would reach targets that are no longer at risk of contagion because they have already adopted the product. Indeed, studies of the spread of infectious diseases show that areas of high network redundancy may actually slow down contagion, reducing the size of the infected population (Badham & Stocker, 2010; Keeling & Eames, 2005).

While this is true for cases in which information about the “product” from contact with one trusted source is enough to trigger adoption, the spread of cultural products may conform to what Centola and Macy (2007) dubbed “complex contagion.” According to them, a contagion process is complex “if its transmission requires an individual to have contact with two or more sources of activation” (Centola & Macy, 2007: 707). Complex contagion is more likely to occur when there is uncertainty about the convenience of adopting the specific product or behavior or when the costs associated with adopting them are not negligible. For such products and behaviors, information from a single source may not suffice to trigger adoption. Rather, potential adopters may require reassurances coming from more than one trusted source in their network neighborhood. The higher the uncertainty or the costs of adoption, the higher the threshold for adoption, and the more important reassurances would be to convert potential adopters into actual ones (see Romero, Meeder, & Kleinberg, 2011 on the effects of complex contagion on the adoption of hashtags on politically controversial topics).

The symbolic nature of cultural products (Bourdieu, 1997) can create considerable uncertainty for the potential adopters, making their adoption a case of complex contagion. This uncertainty, however, may not be equal for all cultural products. Rather, it should increase with the extent to which people may view the adoption of a cultural product as a reaffirmation of a collective identity (Berger & Heath, 2007; Katz-Gerro, 2004; Warde, 1994). Whereas the adoption of a product legitimated by the "community" strengthens one's identification with that community (Childress & Friedkin, 2012), adopting a product that is not aligned with the preferences of relevant others is an indicator that one is not in tune with the tastes and choices of her reference group (Hannan, 2005; Lounsbury & Glynn, 2001; Muller & Peres, 2019; Weick, 1995). At minimum, deviance can create psychological discomfort; at maximum, the deviant consumer could be marginalized from the group (Crane, 1999; Watts, 2002). The more the adoption of a cultural product entails a reaffirmation of the adopter's identity as member of a shared group, the higher the uncertainty she will face before making up her mind about the cultural product. Because a redundant network structure increases the likelihood that potential adopters are exposed to multiple trusted sources of information on a cultural product, the effect of network redundancy on the adoption of cultural products should increase with the extent to which potential consumers view their adoption as a reaffirmation of their identity.

Obtaining adequate measures of the extent to which the potential consumers of a cultural product view the adoption as a reaffirmation of their identity is a daunting task, but certain characteristics of the cultural producers can provide a reliable indicator. Cultural producers can be differentiated using the well-established dichotomy between "generalist" and "specialist" organizations (Freeman & Hannan, 1983; Hannan & Freeman, 1977). Generalists typically target a large and internally differentiated market whose members are less likely to associate the consumption of their products with a reaffirmation of an identity shared with the people who consume their products. The opposite is true for specialist organizations: these cultural producers typically cater to communities with well-defined values and tastes, which are often defined in opposition to the mainstream audiences targeted by the generalists (Bernstein, 1997; Carroll & Swaminathan, 2000). Social movements such as the grass-fed (Weber, Heinze, & DeSoucey, 2008) and the "nouvelle cuisine" (Rao, Monin, & Durand, 2003) provide good examples of such communities. These movements espouse the consumption of products that challenge the codes and values of the mainstream. For members of these communities, the act of consuming certain animal products (e.g., beef from grass-fed animals) or attending certain restaurants becomes a way to reaffirm their values and tastes. Those espousing products or services that may not be viewed as "legitimate" by other community members are likely to be shunned (see Raridon & Mix, 2016 for an example on the grass-fed movement).

We argue that specialist organizations whose cultural products target niche markets are more likely to benefit from redundant online networks than generalists do. Because the adoption of specialist products is likely to be associated with the reaffirmation of shared values, we expect that the redundancy of the online communication network around these products will have a significant positive effect on adoption. The positive effect of network redundancy on adoption should be smaller for cultural products by generalist organizations, because the consumption of generalist's products is less likely to be associated with the reaffirmation of the consumers' identity. The previous discussion results in the following two hypotheses

H1: The adoption of a cultural product increases with the redundancy in the online communication network around that product.

H2: The effect of redundancy in the online communication network on the adoption of cultural products by specialist organizations is stronger than the effect on the adoption of cultural products by generalist organizations.

The previous hypotheses do not consider the content of the messages circulating through the online communication network. This is not uncommon in the contagion literature, because the content of the messages exchanged among potential adopters has been typically unobservable for researchers (but see Berger & Milkman, 2012; Berger & Schwartz, 2011 for recent exceptions). Yet, online social media allows us to overcome this limitation and to consider the effect of the content of the messages reaching potential adopters on their subsequent behavior. We theorize about this effect in the next section of the paper.

The Effect of Network Content

If the adoption of a cultural product were solely driven by the tenor of the messages about that product that circulate within the population at risk, we should expect that adoption should increase with positive evaluations and decrease with negative ones. This simple intuition is confirmed by marketing studies showing that negative publicity hurts product evaluation and sales (e.g., Basuroy, Chatterjee, & Ravid, 2003; Tybout et al., 2019; Wyatt & Badger, 1984). The intuition is also consistent with a baseline contagion model in which the expected attitude of people towards a given product is simply a weighted average of the tenor of the messages to which these individuals are exposed through their social networks, with the strength of the relationships between the individuals providing the weights. The more positive the messages, the higher the likelihood of adopting the product, and vice-versa.

The effect of exposure to contrasting messages on adoption may be less straightforward than this baseline contagion model suggests, however. The reason resides in two distinct mechanisms linking controversy to adoption. On the one hand, and consistent with research on the effect of "social proof" in reducing decision uncertainty (Cialdini, 1984), exposure to both positive and negative messages about a cultural product should reduce the likelihood of adoption, because contrasting messages should contribute to (or at least, fail to reduce) the uncertainty facing the potential adopter. On the other hand, adopting a product requires awareness of the existence of this product (Miner & Mezias, 1996; Strang & Soule, 1998), which could increase with the controversy surrounding that product. This second mechanism suggests that controversy should increase adoption. The marketing literature has demonstrated that the emotional arousal associated with controversy increases product awareness and information sharing (Berger, 2011), bringing it to the attention of people who might not have paid attention to it otherwise. Controversial products and news articles are more likely to be discussed or shared than those that are not (Berger & Milkman, 2012; Luminet, Manstead, & Rimé, 2000). In addition, controversy can trigger spirals of action in social media (Tarrow, 1989; see Vosoughi, Roy, & Aral, 2018 on the spread of true and false news through social media). This is the intuition behind the idea that controversy generates "buzz," increasing the likelihood that people will become aware of the product and engage in conversations about it and potentially search for more information about the product (e.g., Steel 2011). Recent studies confirm that controversy can increase exchanges, especially when social acceptance is not a salient concern for the parties involved in the conversation (Chen & Berger, 2013). This research suggests that we should expect a positive effect of controversy on the adoption of a cultural product.

The presence of two contrasting mechanisms through which controversy about a cultural product can affect the adoption of that product suggests that the net effect of controversy on adoption may depend on the relative salience of each of these mechanisms. If these two opposite mechanisms were equally consequential, we may not be able to observe a noticeable relationship between controversy and adoption. Yet, the effectiveness of these mechanisms may vary depending on the characteristics of the audiences targeted by the cultural product. Again, the characteristics of the cultural producer can provide a good indicator of the characteristics of the audience targeted by its products and offers a solid empirical ground for theorizing about the contingent effects of controversy on adoption.

Cultural products by specialist organizations are more likely to suffer from contrasting opinions since controversy "fosters interpersonal conflicts that jeopardizes social acceptance" (Buss, 1990; Chen & Berger, 2013: 8). Controversy creates further dissent and negative assessments within communities of like-minded people (Coleman, 1990) because social acceptance is more likely to be important among community members. Insofar as specialist cultural producers target consumers that can consider themselves as a virtual community of like-minded individuals, these consumers may view controversial products as a threat to their shared values, making them more likely reject them altogether (Baumeister & Leary, 1995). While controversy may pique their interest about the product, it would also raise the uncertainty associated with the product, making its adoption less likely. The opposite is true for cultural products by generalist organizations. Because these products do not target niche markets of like-minded individuals, they are less likely to suffer from controversy while they can still benefit from the increased awareness and interest of the potential consumers. Therefore, we expect that controversy in online communication networks will decrease the adoption of cultural products by specialist organizations and increase the adoption of cultural products by generalists:

H3: Controversy in the online communication network decreases the adoption of cultural products by specialist organizations.

H4: Controversy in the online communication network increases the adoption of cultural products by generalist organizations.

Data and Methods

We test our hypotheses using data on the sales of electronically published books (eBooks) in the US between 2014 and 2015 and on the Twitter communication networks between people exchanging messages on each of these books online during the same period. Books are good examples of cultural products, whose acceptance and subsequent purchase is affected by the intersubjective judgements of the relevant audiences. Because they are a relatively new phenomenon, eBooks are particularly susceptible to such intersubjective judgements to gain legitimacy, making them ideal subject to test the extent to which their commercial success is affected by the structure and the content on the discussions in online communication networks.

The US publishing industry comprises distinct market players. On one end, the "Big Five" publishers – Penguin Random House, Macmillan, HarperCollins, Hachette, and Simon

& Schuster – are "generalist" core players that mostly target the mass market of mainstream readers and benefit heavily from economies of scale in in-house production, distribution, and marketing. On the other end, there is an unstable periphery of "indie" (i.e., independent) publishers that are not affiliated with the Big Five publishers.[1] In contrast to the Big Five, indie publishers are specialists that serve the minority tastes and espouse "niche" strategies (J. Freeman & Hannan, 1983; Hannan & Freeman, 1977). The eBooks published by indie publishers are typically directed to smaller groups whose members share values and norms that define their identity as members of a virtual community of like-minded individuals (Almandoz, 2012; Marquis & Battilana, 2009). Finally, the publishing industry also hosts a somewhat stable and heterogeneous circle of small and medium publishers (including some well-known academic and university publishers). We keep these publishers in our analysis, but we focus on the contrast between the Big Five and the indie publishers because they best represent the difference between mass and niche producers invoked in our theory.

The sales of eBook are estimated to be about one million paid downloads per day in mid-January 2016 on Amazon. The main data source for eBooks is Author Earnings, an organization that collected and shares information on sales of Amazon books for the purpose of serving indie authors.[2] Author Earnings scraped real-time book rankings on Amazon and uses artificial intelligence algorithms to convert book rankings to daily unit sales. The resulting recency-weighted cumulative daily sales is a much more stable estimate of daily sales than a download of daily sales would have been, as reported by Data Guy (2016). Sales figures are typically within 2% of actual sales for a sample of books that had actual sales data available. The data is collected for a random day in each quarter from 2014 Q1 to 2015 Q3, including

[1] Whereas the "indie publisher" term has multiple meanings for different stakeholders, we adopt the definition that an indie publisher is a publisher not affiliated with any large corporation or conglomerate, as defined by Independent Book Publishers Association.

[2] February 2016 Author Earnings Report: Amazon's Ebook, Print, and Audio Sales. In January 2018, the team behind Authors Earnings launched Book Stats (https://www.bookstats.org) and disabled the original blog.

information on the publisher, genre, reviews, and sales of the books for seven days across the span of seven quarters

Twitter data was obtained from Crimson Hexagon, a company that can provide tweets based on the user-specified keywords. We collected all available tweets (excluding private tweets) sent between the first quarter of 2014 and the third quarter of 2015 that included the Twitter handles @TwitterBooks and @goodreads, which are popular handles for discussing and promoting books. From these messages, we identified all tweets and retweets talking about each of the Amazon eBooks in our sample and matched them to the corresponding book. We used Python to extract the exact author and title matches as well as fuzzy title matches based on the grammatical analysis of the tweet contents (See Appendix A for details). We set the fuzzy match criterion at 86 of 100 since the manual checking of the data using such criterion resulted in a few incorrect matches between tweets and Amazon books.[3] This gives a total of 13,100 tweets and 5,522 retweets during the study period.

We then used the Twitter API to extract followers' information for each of the 10,752 Twitter users responsible for these 18,622 tweets and retweets. The Twitter followership network provides an appropriate representation of the information flows in Twitter, since each tweet or retweet sent by a user is broadcasted across all this user's followers. Because the Twitter API only allows retrieving the list of followers based on the current snapshot of followers, determining the changing followership network throughout our observation period is not straightforward. To do so, we use a snapshot obtained in October 2018 and take advantage of the fact that the Twitter API returns snapshots listing followers in a reverse chronological order, based on the time at which they started following a focal user.

[3] This index is returned by the fuzzy matching routine in Python, which uses the Levenshtein distance between two sequences. Levenshtein distance measures the degree of similarity between two sequences by accounting for the minimum number of edits (insert, delete, and/or substitute elements) that are required for one sequence to exactly match the other sequence (Sankoff & Kruskal, 1983). Then, the fuzzy matching routine in Python computes the Levenshtein distance similarity ratio between the two sequences.

By combining this information with the date in which each of these followers joined the Twitter platform, one can probabilistically infer the followership network of a user in the past (see Vosoughi, Roy, & Aral, 2018 for an example). This method takes advantage of the fact that followers that appear earlier in the focal user list (i.e., those who have started following this user more recently) can only follow this user *after* the date in which other users that appear later in the list (therefore have started following a focal user earlier) have joined the Twitter platform. To check the accuracy of the followers' network inferred from the October 2018 snapshot, we compare it with a snapshot displaying the actual followership network in in November 2017 for 220 randomly selected users. We find that only 4.76 % of the inferred ties are inaccurate. Type I error rate is 1.15% and Type II error rate is 5.04%, mainly due to users that have "unfollowed" a focal user since the time of inferred date till the data collection time.[4]

Once we have established the time-varying followership network for each user that discusses eBooks during the observation period, we identify the Twitter communication network for each eBook in our sample by aggregating time-varying user followership network to book and tweet-month level. For our main analysis, we restrict our sample to the tweets and retweet messages about each eBook sent within one month prior to the random date at which we measure the Amazon sales for that eBook. The size distribution of the resulting eBook Twitter networks is considerably skewed: the 90th percentile for the network size distribution is encompasses 4.06 thousand users, which is well below the maximum of 2.05 million users.

Dependent Variable

Our dependent variable is *daily gross sales* for eBooks on the Amazon platform for a sampled day on each quarter during our observation period, as measured by Authors Earnings.

[4] Descriptively, there is a low likelihood of "unfollowing" Twitter users. Appendix B shows how the number of followers changes over time in our sample between the time of tweet or retweet and the current time. The two examples for @AtriaIndies and @SMoore_Author show that the number of followers drastically increases in the early times of joining Twitter, with subsequent incremental addition of followers without much unfollowing by users.

Because sales figures are skewed (with few books selling a very large number of copies) we use the natural logarithm of book sales in our models, computed after adding one to account for books with no sales in the observed days.

Independent Variables

Network redundancy captures the number of redundant communication paths reaching a user, measured as the number of people the user follows minus one for each node and aggregated across all nodes by book and tweet-month level. We use the natural logarithm of this redundancy measures in our models to compensate for the skewed distribution and minimize the potential influence of outliers on our results.

Controversy captures the variation in sentiment of the messages each node receives about an eBook. Each of the tweets and retweets can have a positive, neutral, or negative tenor, as measured by Crimson Hexagon.[5] Controversy is then calculated as the standard deviation of the tenor for tweets and retweets received by each user, averaged across the book and tweet-month level. We take the natural logarithm of the controversy measures to facilitate the interpretation of the results.

Moderating Variables

We exploit the distinction between publishers targeting the mass market and those focusing on special market niches to test our hypotheses. *Big Five publisher* is an indicator variable for a book published by the Big Five publishers – Penguin Random House, Macmillan, HarperCollins, Hachette, and Simon & Schuster. *Indie publisher* is an indicator variable for indie publishers that are not part of the American Association of Publishers, and it is the omitted

[5] Crimson Hexagon Sentiment Analysis uses a vast set of training posts (over 500,000) that are hand-labelled as positive, negative or neutral in order to predict the tenor of a given post. It calculates the frequency distribution of each word, negated word, emoticon, etc. that is present in documents that are categorized as positive, negative, and neutral. Then, it uses these frequency distributions to construct a model that categorizes each new post to positive, negative, and neutral.

baseline category that comprises of the smallest publishers. We group the remaining eBooks in a *small and medium publisher* category comprising members the American Association of Publishers but significantly smaller (and sometimes more specialized, like university presses) than the Big Five publishers. We dropped self-publishers from the sample since this is a heterogeneous category comprising authors that have accumulated enough fame and resources to publish their own books (e.g., popular authors such as Hugh Howey) along with amateurs whose manuscripts are unlikely to attract the attention of existing publishers.[6]

Social Network Control Variables

Our book level Twitter networks vary considerably along a number of important dimensions such as size, structure, and activity, which may be correlated with our independent variables. We include a number of control variables to account for these differences. *Network size* (logged) to account for the total number of users receiving or sending messages about each book on a given month. Because our measure of *network redundancy* aggregates the number of redundant paths reaching a node at the book (network) level, network redundancy should naturally increase with network size. Thus, controlling for network size is necessary to obtain accurate estimates of the effect of network redundancy on adoption.

Out-degree centralization accounts for the extent to which the Twitter network around a book is dominated by a few nodes followed by many users. We use the network centralization index proposed by Freeman (1978). The index takes the value of 1 for a fully centralized network in which all nodes follow a single node and 0 for a fully decentralized network in which all nodes are equally likely to be followed by other users. The *number of tweets* (logged)

[6] This heterogeneity is reflected on sales figures for self-published books. While books in this category have on average greater daily sales when compared to books by small or medium publishers or indie publishers ($ 28.60 per book per day against $16.85 for small or medium publishers and $ 14.30 for indie publishers), they also have much higher standard deviation of ($ 314.26 dollars for self-publishers against $ 194.33 for small or medium publishers and $ 212.64 for indie publishers).

and *retweets* (logged) on each book and tweet month account for the intensity of the communication about a book. Likewise, because each user may receive multiple messages from each of the people he or she follows, we control for the number of redundant messages that flow through each of these connections. Thus, the *number of redundant messages* is a count of the total number of messages received by each user minus one, aggregated across all nodes by book and tweet-month level. Finally, *no tweet* is an indicator variable set to one for books that do not have any associated tweet during our observation period. All other network variables are set to zero for these observations.

Our models also include a number of control variables to account for the content of the messages on each book. *Tweet tenor* is the average tenor of the messages received by each user, averaged across the book and tweet month level (see Golan & Wanto, 2001; Pollock & Rindova, 2003 for evidence on the effect of tenor on outcomes). Because aggressive marketing efforts using Twitter can affect patterns of adoption (Van den Bulte & Lilien, 2001), our models control for the extent to which the messages received by network members reveal marketing efforts. To this end, we identify Twitter messages that include 'giveaway', 'win', 'chance to win', 'enter to win', 'free copy', 'giving away', 'win a copy', and 'win a signed copy' strings as marketing messages and aggregate them at the user level. The *marketing* variable is the proportion of marketing messages on a given book reaching a user, averaged across all users at the book-month level.

Book attribute control variables

We control for a number of book-level characteristics that may affect sales. *Number of reviews* on Amazon (logged) controls for the fact that the presence of such reviews has a positive effect on book sales (Chevalier & Mayzlin, 2006; Greve & Song, 2017). Likewise, the average star *ratings* for a book controls for the evaluation a book has from prior readers in Amazon, which also affects subsequent sales (Chevalier & Mayzlin, 2006; Mudambi & Schuff, 2010). An indicator variable *unrated* accounts for books without Amazon ratings. We also control for book characteristics such as *sale price*, *days since published* (logged), and genre, of which we include a set of dummy variables for *fiction/literature*, *children's books*, *comics/graphic novels*, *foreign language*, *biography*. *memoir*, and *mystery.*

Estimation Strategy

Because the data has panel structure, the fixed effect estimation is commonly suggested (Wooldridge, 2002: 265-291). However, an FE estimation requires that a book switches publisher type during our observation period or that a publisher switches types, which are relatively rare events and result in small within group-variation. Given this characteristic of our data, and consistent with prior literature (Pathan, 2009), we use a random effect with standard errors clustered by book level. Despite the limitations of the data, we also ran a fixed-effect estimation and found that the pattern of results is consistent with the ones from the random effect models. We report these additional in the robustness analyses of the Results section.

Results

Before proceeding to the presentation of the models testing our four hypotheses, we provide a description of our data to help understand the nature of the US publishing industry. Table 1 displays the descriptive statistics of the variables used in this study using logged values when appropriate to align it with our models. However, we refer to the non-transformed variables in the discussion below to illustrate important characteristics of our sample. The variable *Daily gross sales* is significantly left-skewed prior to logging, with the 90^{th} percentile of gross sales being 59.96 dollars while the maximum is 89,180 dollars. This is consistent with prior findings showing that a limited number of popular books have significantly large sales (Greco, Milliot, & Wharton, 2014).

The distribution of book titles across publisher categories shows that 50.38% the books in our sample are published by indie publishers, which is consistent with studies reporting a long right tail of small specialized publishers (Carroll, 1985; Greco, Milliot, & Wharton, 2014). The remaining books are evenly split between the Big Five (24.78 %) and the small and medium publishers (24.85 %). However, average sales per book are the lowest for indie publishers at $14.30, increasing to $16.85 for small and medium publishers and to $101.78 dollars for the Big Five. Publishers also differ in terms of genres. General interest genres such as “non-fiction” and “fiction” are significantly more common among the Big Five than among indies. Conversely, niche genres like “comic and graphic novels” or “adult romance and fantasy” are significantly more likely to be represented among indie publishers (Data Guy, 2016). Finally, 98 percent of the books in the sample do not have any tweets associated with them in the month prior to the randomly selected day in which we measure sales, as indicated by the value of the *no tweet* dummy variable.

=== Insert Table 1 about here ===

The differences among publisher categories are also apparent in the structure of the Twitter networks associated with the books in our sample. Figure 1 shows graphic representations of these networks. The top three graphs display the Twitter networks of three books published by the Big Five, whereas the bottom three graphs display the networks for titles by indie publishers. In both cases, the books were selected to minimize the difference between their actual and expected sales (as predicted by the Twitter network structure for each category) to illustrate the differences in these networks. The networks corresponding to books published by the Big Five are typically bigger and centralized around a few users, which often are the publishers themselves. Indeed, the users with the largest outdegree in in Figure 1A correspond to Little Brown (part of Hachette), Simon & Schuster, and Berkeley Romance (part of Penguin Random House) respectively, whereas the nodes with the second largest outdegree are the authors, which are encouraged by the publishers to have an active presence in social media.

=== Insert Figure 1 about here ===

The structural pattern is different for the indie books displayed in Figure 1B. Although some of these networks may also be centralized around active players (e.g., the bottom right panel), this is less common and the outreach is much smaller, as illustrated by the graphs displayed in the other two bottom panels. Moreover, the most central users in these graphs are community members with an interest in the genre rather than publishers or authors. These differences are reflected in the statistics in our sample. Considering books that have at least one tweet, the average network size is 5,015 users for Big Five books and 3,642 for indie books. Likewise, the node with the largest number of active followers has 4,390 followers for Big Five books and 3,011 users for indies. These figures are consistent with the idea that there are substantial differences in network size and between the Twitter networks on books published by the large Big Five publishers and those associated with indie books.

Table 2 presents the results of our analysis. Model 1 is the baseline model with all of the control variables and independent variables. Model 2 includes interactions terms for *network redundancy* and Model 3 does the same for *controversy,* the two key independent variables in the analysis. Finally, Model 4 includes all the interactions in the analysis. For simplicity, we will discuss the results focusing on Model 4 when appropriate, but the pattern is consistent with the ones reported in the partial models. Because the effect of network redundancy may be affected by network size and size varies substantially across publisher type, we include interactions between this variable and type of publisher in models that include the *network redundancy* variable. This allows us to interpret the effects of network redundancy by publisher type net of the confounding effect of network size.

Hypothesis 1 states that the adoption of a cultural product should increase with the redundancy in the online communication network around that product. The insignificant coefficient for *network redundancy* in Model 1 does not provide support for this hypothesis. However, Model 4 (as well as Model 2) shows that the net effect of *network redundancy* on book sales is positive and significant for indie publishers ($\beta = 0.124$, $p < .01$. Holding other variables constant, a 10 percent increase in network redundancy results in a 1.24 percent increase in book sales for indies. Hypothesis 2 is supported, as the slope for Big Five publishers is significantly different from the baseline ($\beta = -0.150$, $p < .01$). The difference is also significative for small and medium publishers ($\beta = -0.172$, $p < .01$). Yet, the net effects of network redundancy for these two categories of publishers are negligible and not significant. These results suggest that, unlike indies, the Big Five publishers and small or medium publishers do not reap benefits from network redundancy. Figure 2 reports graphic illustration of the average adjusted predictions for the effect of network redundancy on book sales by indie and Big Five publisher types.

The negative effect of network size on the sales of indie books ($\beta = -0.086$, $p < .01$) deserves special attention. Given that network size captures the number of Twitter users receiving or sending messages about the book, one would expect a net positive relationship. Indeed, this is the case for books by the Big Five publishers ($\beta = 0.053$, $p < .01$). Additional analyses reveal that the negative effect of network size for indie titles is in fact driven by outdegree network centralization, which is strongly correlated with size in our data ($r = 0.89$). When an interaction between centralization and publisher type is included in the model, the net effect of network size on indie sales becomes positive ($\beta = 0.18$, $p <0.05$), without altering our main results. Centralization, in turn, has a strong negative effect on sales of indie books ($\beta = -0.720$, $p < 0.01$). This suggest that, unlike books by the Big Five publishers, attempts to reach a larger audience through highly connected nodes may backfire for indie books, as users engaged in conversations about those books may view this as a violation of the community norms on the nature of their Twitter exchanges.

Because our models control for the effect of network size by publisher type, we know that the effects of redundancy on sales are not driven by the size of the network but by its structure.[7] We further examined our interpretation by looking at the count of nodes with redundant ties in the network, because similar levels of network redundancy can be obtained if redundant ties are concentrated on a subset of the nodes in the network of if these are spread evenly across the network. Holding network size constant, the number of nodes with at least one redundant tie have a significant positive net effect on book sales for indies and no effect for any of the other two categories of publishers, which is a pattern that is similar to the one reported in Model 4. Further increases of redundancy at the node level do not increase (and may even hurt) indie sales. This additional analysis confirms that redundancy at the node level is responsible for our findings, which is consistent with the complex contagion model. The greater number of nodes that receive reassurances from two independent sources is related to the higher the sales of indie books.

The fact that the expected benefits of network redundancy on the adoption of cultural products only become apparent for books targeted at niche markets is consistent with the idea that the adoption of such books can be modeled as a complex contagion model. Yet, we still expected that the benefits of redundancy for books targeted at the mass market would be smaller than those for books targeted at niche markets, but not negligible. One possible explanation is that redundancy may have two opposite effects on book sales. While redundancy contributes to the legitimacy of the book before potential buyers, it also restricts the circle of people who hear about the book to one or more "closed" network neighborhoods, hindering

[7] Because our measure of network redundancy aggregates the number of redundant paths reaching a node at the book (network) level, network redundancy naturally increases with network size (r = 0.56, Table 1). Thus, controlling for network size is necessary to obtain accurate estimates of the effect of network redundancy on adoption.

the spread of the message to a broader audience. When the need for reassurance from trusted sources is high, the positive effects of redundancy on the likelihood of adoption outweigh its costs. Conversely, when the need for reassurance is low, the cost of redundancy may offset (or even outweigh) its benefits, which can result in a negligible or even negative effect of redundancy on adoption. Because people are likely to buy the book only once, additional redundant messages may reach individuals who are no longer at risk of buying the book, in a process that is akin to the spread of diseases in tight communities (Granovetter, 1973; Onnela et al., 2007; Watts, 1999). The conjecture that the costs of redundancy may offset its benefits for generalists is consistent with the fact that, unlike specialists, generalists do benefit from network size, as reported before ($\beta = 0.053$, $p < .01$). This suggest that the adoption of cultural products by generalist producers may correspond more to a simple contagion process, whereas those of specialist producers conforms to the complex contagion model.

=== Insert Table 2 about here ===

Finally, it is worth noting that the positive effect of network redundancy on the sales of indie books is driven by messages from multiple sources, and not by the multiplicity of messages from the same source. Two facts back up this claim. First, the effect of network redundancy is net of the effect of redundant messages, which is a control in the model. Second, an additional analysis in which we re-estimated Model 2 with interaction terms between the number of redundant messages and publisher type reveal that the net effect of the number of redundant messages on sales is not significant for any of the publishers, whereas the effects of network redundancy (i.e., the number of active paths reaching the target) remains. These results suggest that, consistent with the complex contagion model, it is the number of redundant *sources* (and not just the number of messages from the same sources) what drives adoption.

The interactions between *controversy* and publisher type test the hypotheses on the contingent effect of controversy on adoption. The effect of controversy is negative and significant for indie publishers (β = -1.202, p < .05), which provides support for H3. The estimate suggests that a 10 percent increase in controversy decreases indie book sales by 12.02 percent. The opposite is true for the Big Five publishers. For this category, the effect of controversy on book sales is significantly positive (β = 2.974, $p < .01$) and so is the net effect of controversy on Big Five book sales: a 10 percent increase in controversy increases Big Five book sales by 17.72 percent. The effect of controversy is also positive and significant for the small and medium publishers (our residual category), but the heterogeneity of this group prevents us from making claims about this finding. The effects of controversy on book sales are slightly stronger in the partial Model 3, but the magnitudes and confidence intervals are comparable with those of Model 4. Figure 3 reports graphic illustration of the average adjusted predictions for the effect of controversy on book sales by indie and Big Five publisher types.

Taken together, our results provide support for H3 and H4: controversy hurts adoption for products targeted at groups that are likely to view adoption as a reaffirmation of a common identity, but it increases adoption when this is not the case, presumably due to the increase in "buzz" and product awareness associated with controversy. This conjecture is confirmed by an additional analysis revealing that, after controlling for book attribute control variables, average tenor, and marketing content, a 10 percent increase in controversy increases the number of tweets and retweets on a book by 8.9 percent ($p < .01$). For indies, however, the potential benefits of this additional buzz are likely to be outweighed by the increase in uncertainty facing potential adopters.

=== Insert Figure 2 and 3 about here ===

Robustness Tests

We conduct several tests to evaluate the robustness our findings. First, we used an alternative measure of network redundancy that considers the local structure of a user's Twitter network. Specifically, we counted the number of closed triads in each node's neighborhood (logged), defined by the users followed by this node. A triad is "closed" if one or two users followed by the focal node follow the other user. From a theoretical viewpoint, this alternative measure considers the potential effect of indirect communication paths among the people sending messages to the focal user on the intensity and consistency of the messages reaching this user. The two measures are highly correlated in our sample (r = 0.890, p < 0.01), largely due to the fact that they are both aggregates at the network level. The estimates of the main effect of network redundancy on indie sales obtained by replacing our original measure with the triad-based measure are similar (β = 0.133, for the number of triads vs. β = 0.124 for our original measure, p < .01), whereas the net effects for the other two publisher categories remain negligible. These results, added to the fact that the triad count measure does not change the model fit (R^2 = 0.4263, compared to original R^2 = 0.4264), suggest that the additional effect of triadic closure on redundancy is marginal in our data.

Second, we measured controversy as the overall variance of message tenor directly at the book and tweet-month level, instead of computing it first at the user level and then aggregating it at the book and tweet-month level. While our original measure tries to capture the level of controversy reaching each individual user, the alternative measure focuses on the overall controversy about the book. Although we are able to reproduce our results with this alternative measure, the effect sizes are smaller. A 10 percent increase in book-level controversy results in a 5.20 percent decrease in sales for indie published books instead of the 12.02 percent decrease when using the original node-level measure. Likewise, a 10 percent increase in controversy results in a 4.83 percent increase in sales for the Big Five published books, against 17.72 percent increase when using the node-based measure. The significant decrease in effect sizes when using the overall book-level controversy measure suggests that the ego-level variation in sentiment best captures the effect of controversy on individual decisions, which aggregate to total book sales.

Third, certain book authors may enjoy levels of prestige and legitimacy comparable or even bigger than those of their publishers. To account for this, we computed a measure of author prominence as the total number of tweets and retweets talking about each author (See Appendix A for details on how we extract the author names from tweets). This measure, however, is strongly correlated to the aggregate measure of the total number of tweets and retweets talking about a specific book or an author (r =0.855, $p<0.01$), because a book and its author are often co-mentioned in the same tweet. Although author prominence has the expected positive effect on sales, the inclusion of this variable does not alter our results or increase model fit. These findings suggest that our results are not affected by author's prominence.

Fourth, whereas we control for the fact that the majority of the books did not have any tweeting activity by identifying such books with an indicator variable in our models, it is possible that our results might have been affected by the large number of books without tweets. To examine this possibility, we run our models on the subsample of books that have at least one associated tweet. To account for the possibility that these books are a non-representative sample of the population, we first estimate book sales for the population as a function of all our non-Twitter independent variables (excluding publisher type) and then used the predicted value of book sales as a regressor in our models on the restricted sample. In addition of predicted sales, these models include our Twitter variables and the publisher type moderators. The pattern of results on this restricted sample is largely consistent with the one reported before, albeit the effect sizes are often smaller. Network redundancy has a marginally significantly positive effect on book sales for indie publishers in this reduced sample ($\beta = 0.056$, $p < .10$) whereas the net effect of redundancy on sales remains insignificant for the other two publisher categories, which is consistent with our main results. The positive effect of controversy for the Big Five publishers is still present ($\beta =1.045$, $p < 0.01$), although the negative effect for indies becomes statistically insignificant ($\beta =-0.610$, $p = 0.19$). It is important to note that some of the variance associated with the Twitter activity in our main models was likely absorbed by non-Twitter correlates in the model used to generate our sales estimates for the whole sample, which would account for the weakening of the results. The fact that we were able to reproduce the pattern of results despite this limitation provides additional credibility to our findings.

Fifth, we recomputed our independent variables using messages issued during the last two, three and four months before the date at which we measure book. The effect of network redundancy on indie book sales is stronger when this variable is computed over the messages sent during the prior month, which is the window used in our main analysis (β =0. 124, $p < 0.01$). The effect then gradually decreases as the window widens (β =0. 072, $p < 0.01$ for a four month window). These estimates show that the effect for network redundancy for tweets in the most recent month is 1.5 times larger than the effect of network redundancy computed over the last four months. The decrease is more striking for controversy. In this case, all coefficients are not statistically significant beyond the one-month window. These results support our choice of a one-month window to estimate the effects of the structure and the content of Twitter messages on book sales and suggest that those effects are largely driven by the most recent exchanges among Twitter users.

Finally, we examined the possibility that that some unobserved book characteristics could be simultaneously driving Twitter communication and book sales, raising endogeneity concerns. We sought to address these concerns in two different ways. First, we used a book-fixed effects estimation to account for any unobserved book characteristics. This estimation requires that a book would change publisher during the sampled period (e.g., a title that was originally issued by an indie publisher is reissued by a Big Five publisher) or that a publisher changes category (e.g., an indie publisher is bought by one of the Big Five), which results in a substantial reduction in sample size and we end up with an average of 1.7 observations per group. Despite this limitation, we were able to reproduce the pattern of results with the fixed-effect estimation. The effect of network redundancy on indie book sales is positive and marginally significant ($\beta = 0.080$, $p < 0.05$) whereas the net effect fort the Big Five is insignificant ($\beta = 0.031$, $p = 0.219$), which is consistent with the results from Model 4. Likewise, the effect of controversy on indie book sales is negative ($\beta = -0.861$, $p < 0.05$). The only effect that differs from our main models is that of controversy on Big Five books: although still positive, the net effect of controversy on sales is no longer significant ($\beta = 0.324$, $p = 0.18$), which may simply result from limited within-group variability in the fixed effects model.

A second possibility is that both current book sales and the Twitter activity in the prior month are driven by prior sales, which may render the relationship between our independent and dependent variables spurious. Although the non-stationary nature of our data prevents the inclusion of the lagged dependent variable (Keele & Kelly, 2006), we can still examine whether the conditions that could lead to a spurious relationship are present in our data—that is, that network redundancy and controversy are a function of prior book sales. To this end, we estimated these two independent variables as a function of prior book sales and all our control variables. The effect of prior sales on network redundancy is negligible ($\beta = -0.005$, $p < 0.01$), whereas the effect on controversy is not statistically significant ($\beta = -7.47e\text{-}06$, $p = 0.484$). In other words, there is no evidence of a significant relationship between prior book sales and our independent variables. If prior book sales do not have a significant effect on network redundancy or on controversy, the observed effects of these variables on subsequent book sales cannot be attributed to the spurious influence of prior sales.

Discussion

This paper shows that the effect of online communication networks on the adoption of cultural products is contingent upon the characteristics of the markets targeted by the cultural producers. "Specialist" cultural producers target niche markets, whose members are are often part of a community defined by the consumption of a particular type of cultural products.

Consistent with the predictions derived from models of complex contagion (Centola & Macy, 2007), these specialist producers benefit from the reassurance provided by redundant communication paths reaching potential adopters. Communication redundancy, however, does not benefit "generalist" producers targeting a large, undifferentiated collection of individuals that simply share information and comments about their products but do not necessarily view each other as sharing the same values or as members of a virtual community. The differences across these two types of cultural producers is more pronounced with respect to the content of the online messages. Because the potential adopters of a specialist product are more sensitive to the uncertainty created by conflicting online messages around the product within communities, controversy decreases adoption for specialist cultural products. The opposite is true for generalists, which benefit from the increasing buzz generated by controversy in the online messages around the product.

Classical contagion studies have emphasized the importance of network reach over network redundancy in explaining diffusion (e.g., Granovetter, 1978), whereas more recent models vindicate the importance of redundancy as a way to diminish the uncertainty potential adopters face when making their decision (Centola & Macy, 2007). Our findings suggest that the pertinence of these models in explaining the adoption of cultural products may depend on the extent to which the adoption entails identity claims. The more the adoption of a cultural product can be viewed as the enactment of values shared with other members of a virtual "community," the more cultural producers will benefit from network redundancy and from the absence of controversy, which is consistent with the predictions of the complex contagion model. The opposite is true for cultural products whose adoption does not entail strong identity claims. The spread of such products is more likely to follow the "simple contagion" model protracted by classical diffusion studies. The multiple and consistent sources of information are less necessary to reassure the potential adopters. At the same time, the redundant sources may undermine the reach of the information and its capacity to attract the interest of those adopters.

Our findings also highlight the importance of the content of the messages that circulate through the network, above and beyond the effect of network structure in explaining adoption. Because the content of the messages circulating through network tie was not easy to observe, scholars interested in the effect of network content on outcomes were largely limited to use network composition (i.e. the stock of knowledge controlled by an actor's contacts) as a proxy for the information exchanges between the focal actor and those contacts (e.g., Podolny & Baron, 1997; Rodan & Galunic, 2004). The "breadcrumbs" left behind in online social networks allow us to overcome these limitations. The marketing literature has made some important advances in studying message content (Berger & Milkman, 2012; Berger & Schwartz, 2011), but scholars are just starting to examine how the interplay between network structure and content affect outcomes. The contingent effect of network structures on outcomes is well established in the literature that examines intra and inter-organizational networks (e.g., Berger and Milkman 2012; Berger and Schwartz 2011; Burt, 1997; Gargiulo, Ertug, & Galunic, 2009). Our findings suggest that the effect of network content can also vary with the nature of the actors whose behavior may be affected by this content. Future research can benefit from the ability to observe both the structure and the content of the communication in online social media to explore different dimensions of network content as well as their interactions with network structures.

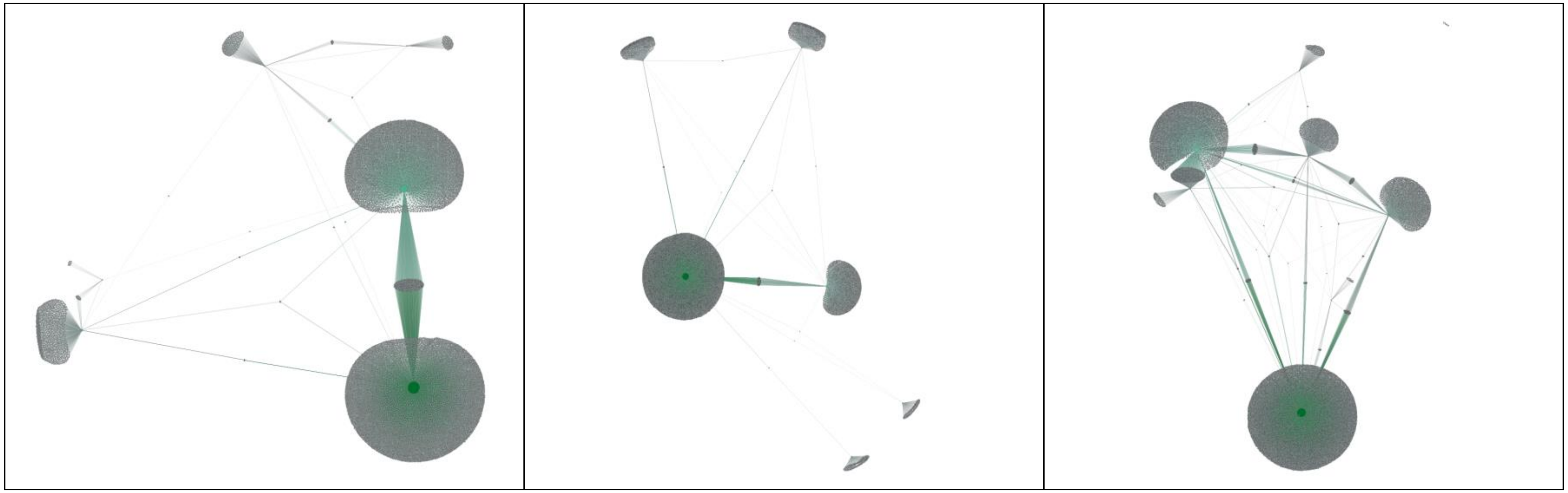
In addition to contributing to our understanding of the effects of network structure and content on the adoption of cultural products, our findings also have practical implications for the digital strategies of cultural producers that seek to leverage social media platforms. Given the opportunities to reach a large number of potential consumers at negligible marginal cost, one might expect that cultural producers should take advantage of online platforms to reach as many potential consumers as possible. Indeed, generalists like the Big Five publisher houses can leverage their centralized marketing efforts and reach large numbers of people without being concerned with the potential for divergent views about their product, as the buzz generated by divergent views actually benefits the adoption of their products. Our findings, however, suggest that such strategy may backfire for specialist producers such as indie publishers, whose products are adopted as part of the reaffirmation of shared values among community members. Rather, these specialists can gain competitive advantage by leveraging redundant ties in niche communities while seeking to minimize controversy. Because the adoption of specialist products depend on community mobilization, the successful strategy for specialist cultural producers resemble more those of grassroot social movements than those of their mainstream industry peers (Rao et al., 2003; Weber et al., 2008).

Our study also sheds light on the effects of social media platforms on the competitive dynamics of cultural industries. While existing research converges in arguing that digital platforms like social media have disrupted the competitive dynamics and the relative standing of firms within many industries, there has been conflicting evidence on the direction of change. Some studies have claimed that powerful firms can leverage online platforms to further enhance their market power (Gawer & Cusumano, 2014), whereas others suggest that peripheral players can take advantage of those platforms (Gans & Stern, 2003) to increase the competitive pressures from peripheral players (Benner & Waldfogel, 2016; Carroll & Swaminathan, 2000; Greve & Song, 2017) and generate grassroot movements that can change the dynamics of the industry (Kotha & Basu, 2011). Our study of book publishers suggests that the effect of social media on the competitive dynamics of the industry may be more complicated than what the current literature suggests. Whereas social media offer both generalist and specialist "niche" cultural producers a relatively cheap means to reach more potential consumers, the ability to leverage this means to increase the adoption of their products varies substantially by type of producer. While generalists that target an undifferentiated mass market may use their resources to reach large numbers of potential consumers at a negligible marginal cost and even benefit from the buzz, this is not the case for specialists. Because adopting their products is more likely to entail a reaffirmation of certain values and identity, the success of their social media strategy depends more on the messages that circulate *among* community members than on the messages that go from the publisher to those members. Therefore, while social media platforms can provide opportunities to change industries from the periphery, it may be harder for specialist cultural producers to take full advantage of such opportunities due to the difficulty in community coordination.

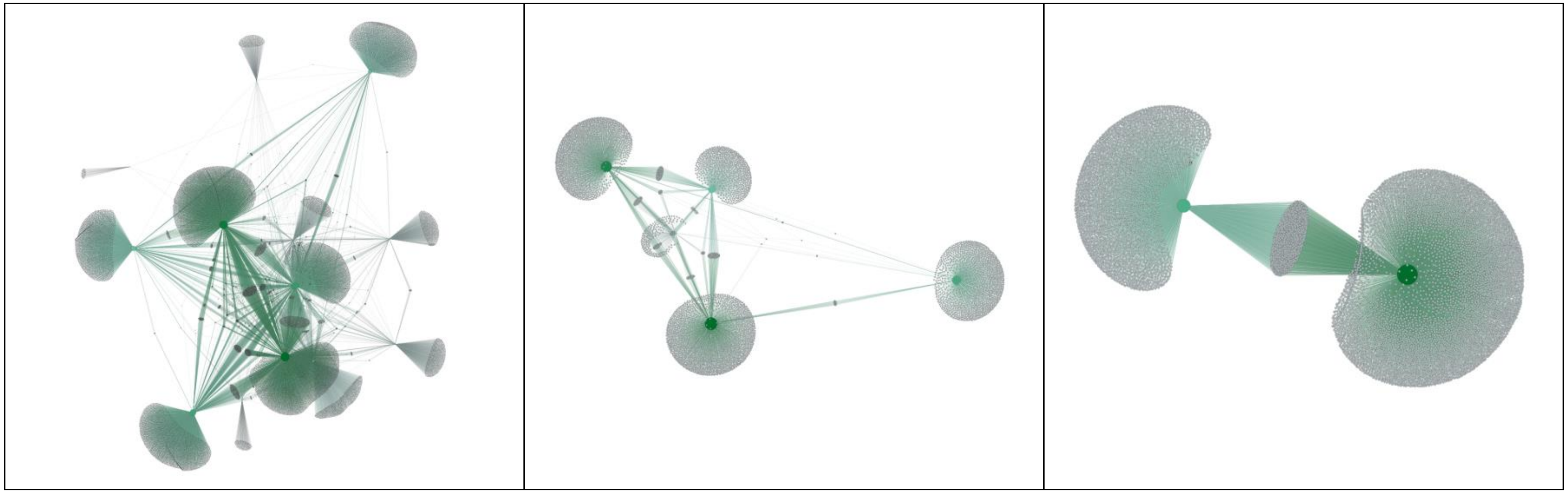
Despite the richness of our data that allows us to observe the structure and the content of the online communication network around a book as well as their effect on the online sales of that book, we cannot observe individual adoption. That is, we do not know whether or when a particular Twitter user has bought the book, nor the instances in which buyers learned about the book through other means. We sought to alleviate this limitation of our data through a clear specification of the theoretical mechanisms responsible for adoption at the individual actor level. In addition, we measured our independent variables (network redundancy and controversy) at the actor level and then aggregate these measures at the book level and matched them with sales figures. In an ideal scenario, we would have time-stamped data on individual book purchases and would have been able to model the likelihood that an individual would purchase the book as a function of the structure and content of his or her network neighborhood. Because these data were not available, we used aggregate adoption figures and matched them with our network variables, in a way that is akin to the approach adopted by traditional epidemiology studies that cannot access detailed data on infected individuals, albeit with the difference that we do have data on networks of information flow between individuals. While we acknowledge the limitations of our approach, we argue that such limitations could become significant if book sales were driven by unobserved factors that are correlated with the structure and the content of the online communication network, or if the bulk of the sales were accounted for by people who were not exposed to social media. We do not have any reason to believe that this is the case. Therefore, we believe our study provides an accurate—albeit imperfect—depiction of the processes through which individuals decide to adopt cultural products.

By highlighting the contingent effects of network content on the adoption of cultural products, our study advances our understanding of the effects of the structure and content of the messages that circulate through online social network on the adoption of cultural products. From a structural viewpoint, the extent to which the adoption of cultural products follows a "simple" or a "complex" contagion model depends on the extent to which the adoption can be viewed as the enactment of certain shared values by members of a virtual community. The more this is the case, the more the adoption conforms to a complex contagion model, where network redundancy—and not simply reach—is important to trigger adoption. The results for network controversy, however, are not always consistent with existing network models of contagion. These models have previously assumed that the expected response of a given actor is a weighted average of the responses of the actor's neighborhood. Therefore, consistently positive messages should increase adoption and the opposite should be true for consistently negative ones. The effect of controversy, defined as a mix of positive and negative views, should fall somewhere in between. Yet, our findings suggest that controversy can increase or decrease adoption depending on the nature of the type of cultural producer, with mainstream producers benefiting from it and nice "specialist" producers suffering from it. These findings open the way for further research on the contingent effect of network content on outcomes across different types of contexts, which should expand our understanding of how network structure and content affect adoption.

Figure 1. Network structure for Big Five published books (1A) and indie published books (1B)



1A. Big Five Publishers



1B. Indie Publishers

Table 1: Descriptive statistics and correlations

	Variables	Mean	SD	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
1	Daily gross sales (ln)	0.98	1.78	1.00										
2	Network redundancy (ln)	0.01	0.23	0.07	1.00									
3	Controversy (ln)	0.00007	0.01	0.03	0.06	1.00								
4	Number of redundant messages (ln)	0.01	0.22	0.04	0.43	0.28	1.00							
5	Tweet tenor	0.004	0.07	0.07	0.08	0.08	0.09	1.00						
6	Marketing	0.001	0.03	0.05	0.42	0.04	0.19	0.04	1.00					
7	Network size (ln)	0.09	0.75	0.16	0.56	0.09	0.32	0.40	0.36	1.00				
8	Outdegree centralization	0.01	0.12	0.17	0.27	0.10	0.19	0.42	0.24	0.89	1.00			
9	Number of original tweets (ln)	0.02	0.13	0.16	0.47	0.12	0.31	0.40	0.23	0.91	0.85	1.00		
10	Number of retweets (ln)	0.002	0.05	0.05	0.84	0.08	0.47	0.08	0.41	0.44	0.20	0.29	1.00	
11	No tweet	0.98	0.13	-0.17	-0.36	-0.10	-0.22	-0.43	-0.26	-0.92	-0.98	-0.92	-0.27	1.00
12	Total reviews (ln)	2.54	1.79	0.53	0.05	0.02	0.03	0.07	0.03	0.14	0.15	0.15	0.04	-0.15
13	Average rating	3.62	1.7	0.18	0.01	0.01	0.01	0.01	0.01	0.02	0.03	0.02	0.01	-0.03
14	Unrated	0.17	0.37	-0.18	-0.01	-0.00	-0.01	-0.02	-0.01	-0.03	-0.03	-0.03	-0.01	0.03
15	Sale price (ln)	2.34	0.75	-0.12	-0.01	-0.00	-0.01	-0.02	-0.01	-0.04	-0.04	-0.04	-0.01	0.04
16	Days since published (ln)	6.6	1.5	-0.17	-0.01	0.00	-0.00	0.00	-0.01	-0.00	-0.00	0.00	-0.02	0.00
17	Big Five publisher	0.25	0.43	0.28	0.05	0.01	0.03	0.05	0.03	0.12	0.12	0.12	0.03	-0.13
18	Small or medium publisher	0.25	0.43	-0.10	-0.02	-0.00	-0.01	-0.02	-0.01	-0.05	-0.05	-0.05	-0.01	0.05
19	Indie publisher	0.5	0.5	-0.16	-0.02	-0.01	-0.01	-0.03	-0.02	-0.06	-0.07	-0.07	-0.02	0.07
20	Fiction/literature	0.16	0.37	0.28	0.07	0.02	0.03	0.07	0.05	0.16	0.17	0.16	0.04	-0.17
21	Children’s books	0.05	0.22	-0.02	-0.00	0.01	0.00	0.00	-0.00	-0.01	-0.00	-0.00	0.00	0.00
22	Comics/graphic novels	0.02	0.15	-0.01	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	0.00
23	Foreign language	0.03	0.17	-0.08	-0.00	0.00	0.00	-0.00	-0.00	0.00	0.00	0.00	-0.00	-0.00
24	Memoir	0.09	0.28	0.04	-0.01	-0.00	-0.00	-0.00	-0.00	-0.01	-0.01	-0.01	-0.00	0.01
25	Mystery	0.03	0.16	0.20	0.04	0.03	0.04	0.02	0.03	0.09	0.09	0.09	0.04	-0.09

Table 1: Descriptive statistics and correlations (cont'd)

	Variables	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)
13	Average rating	1.00												
14	Unrated	-0.95	1.00											
15	Sale price (ln)	0.02	-0.01	1.00										
16	Days since published (ln)	0.22	-0.24	0.13	1.00									
17	Big Five publisher	0.13	-0.15	-0.03	0.03	1.00								
18	Small or medium publisher	0.04	-0.04	0.31	0.08	-0.33	1.00							
19	Indie publisher	-0.15	0.16	-0.24	-0.09	-0.58	-0.58	1.00						
20	Fiction/literature	0.01	-0.03	-0.25	-0.10	0.14	-0.13	-0.01	1.00					
21	Children's books	0.02	0.00	-0.15	-0.04	0.07	-0.02	-0.04	0.01	1.00				
22	Comics/graphic novels	-0.02	0.02	-0.07	-0.07	-0.05	-0.04	0.08	-0.03	0.00	1.00			
23	Foreign language	-0.24	0.25	-0.09	-0.03	-0.07	-0.09	0.13	0.05	-0.00	0.02	1.00		
24	Memoir	0.06	-0.07	-0.01	0.01	0.14	-0.04	-0.09	-0.09	-0.04	-0.04	-0.02	1.00	
25	Mystery	0.01	-0.02	-0.08	-0.04	0.12	-0.07	-0.04	0.30	-0.03	0.02	0.04	-0.05	1.00

N=498,736 except for network variables with N=498,586

Table 2: Main results

	Model 1	Model 2	Model 3	Model 4
Total reviews (ln)	0.581**	0.581**	0.581**	0.581**
	(0.002)	(0.002)	(0.002)	(0.002)
Average rating	0.075**	0.075**	0.075**	0.075**
	(0.004)	(0.004)	(0.004)	(0.004)
Unrated	0.959**	0.959**	0.959**	0.959**
	(0.017)	(0.017)	(0.017)	(0.017)
Big Five publisher	0.249**	0.238**	0.249**	0.238**
	(0.007)	(0.007)	(0.007)	(0.007)
Small or medium publisher	0.101**	0.097**	0.101**	0.097**
	(0.006)	(0.006)	(0.006)	(0.006)
Sale price	-0.061**	-0.062**	-0.061**	-0.062**
	(0.003)	(0.003)	(0.003)	(0.003)
Days since published	-0.316**	-0.316**	-0.316**	-0.316**
	(0.002)	(0.002)	(0.002)	(0.002)
No tweet	-0.250*	-0.352**	-0.248*	-0.347**
	(0.107)	(0.108)	(0.107)	(0.108)
Fiction/literature	0.567**	0.568**	0.567**	0.568**
	(0.007)	(0.007)	(0.007)	(0.007)
Children’s books	-0.244**	-0.243**	-0.243**	-0.243**
	(0.008)	(0.008)	(0.008)	(0.008)
Comics/graphic novels	-0.022	-0.021	-0.022	-0.021
	(0.014)	(0.014)	(0.014)	(0.014)
Foreign language	-0.125**	-0.121**	-0.124**	-0.121**
	(0.007)	(0.007)	(0.007)	(0.007)
Memoir	-0.086**	-0.084**	-0.086**	-0.084**
	(0.009)	(0.009)	(0.009)	(0.009)
Mystery	0.693**	0.694**	0.692**	0.694**
	(0.016)	(0.016)	(0.016)	(0.016)
Number of original tweets (ln)	0.057	-0.021	0.057	-0.017
	(0.061)	(0.062)	(0.060)	(0.062)
Number of retweets (ln)	0.023	-0.009	0.027	-0.005
	(0.086)	(0.087)	(0.085)	(0.087)
Tweet tenor	0.022	0.018	0.023	0.019
	(0.033)	(0.032)	(0.033)	(0.032)
Marketing	0.130+	0.123	0.133+	0.125+
	(0.075)	(0.075)	(0.075)	(0.075)
Network size (ln)	0.014	-0.088**	0.014	-0.086**
	(0.009)	(0.011)	(0.009)	(0.011)
Outdegree centralization	-0.063	-0.112	-0.062	-0.110
	(0.093)	(0.093)	(0.093)	(0.093)
Number of redundant messages (ln)	0.017	0.021+	0.016	0.020+
	(0.012)	(0.012)	(0.012)	(0.012)

Network redundancy (ln)	0.006	0.124**	0.005	0.124**
	(0.022)	(0.029)	(0.022)	(0.029)
Controversy (ln)	0.672*	0.704*	-2.024**	-1.202*
	(0.323)	(0.305)	(0.522)	(0.524)
Big Five publisher X network redundancy (ln)		-0.148**		-0.150**
		(0.028)		(0.027)
Big Five publisher X network size (ln)		0.141**		0.139**
		(0.008)		(0.008)
Small or medium publisher X network redundancy (ln)		-0.169**		-0.172**
		(0.046)		(0.045)
Small or medium publisher X network size (ln)		0.123**		0.123**
		(0.014)		(0.014)
Big Five publisher X controversy (ln)			4.018**	2.974**
			(0.639)	(0.643)
Small or medium publisher X controversy (ln)			1.599**	0.684
			(0.559)	(0.567)
_cons	1.387**	1.494**	1.385**	1.489**
	(0.110)	(0.110)	(0.110)	(0.110)
Ll_0	.	.	.	.
Ll	.	.	.	.
N	498,586	498,586	498,586	498,586
Df_M	23	27	25	29

Figure 2. Effect sizes for *network redundancy* with respect to publisher category

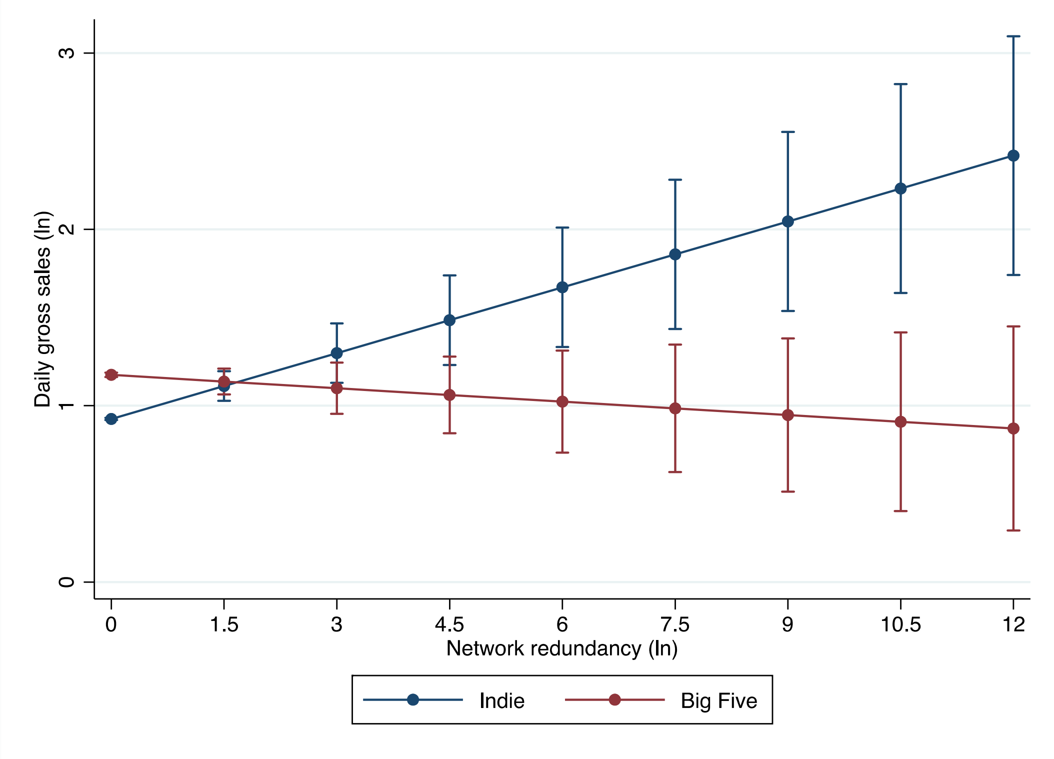
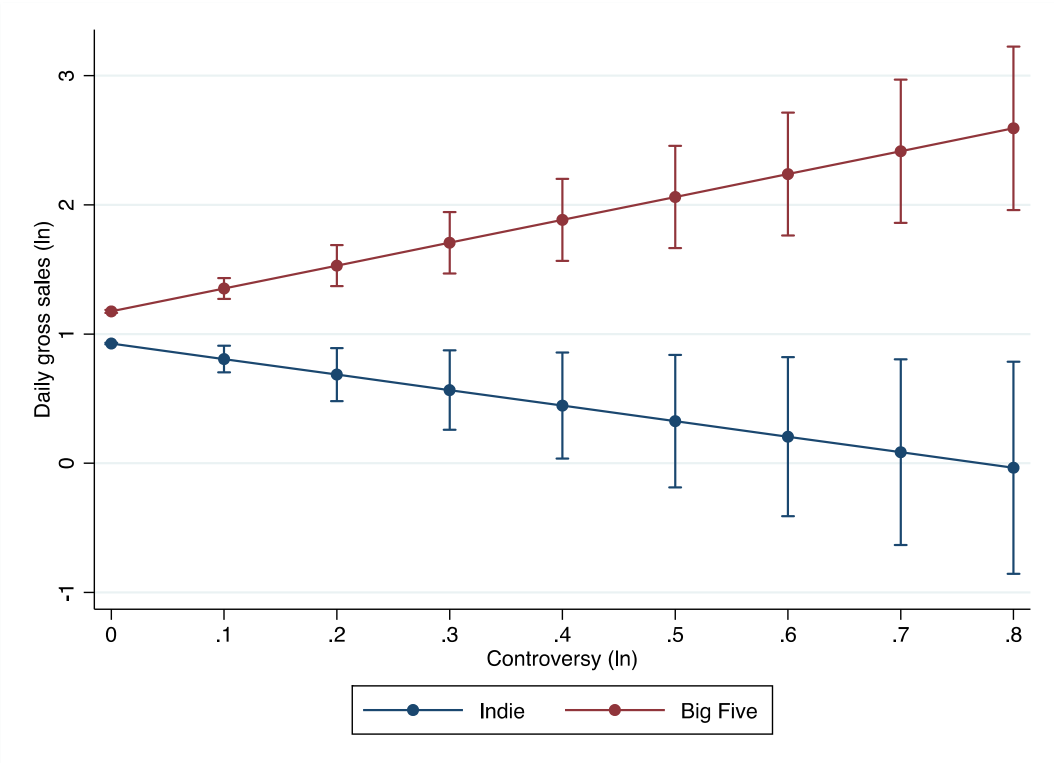


Figure 3. Effect sizes for *controversy* with respect to publisher category



Appendix A. Matching Twitter Contents to Books

Matching Twitter contents to books involves *1) cleaning up the contents, 2) extracting author names from the text, 3) extracting book titles from the text*, 4) *matching exact author and book titles*, and 5) *fuzzy matching book titles*.

1. Cleaning up the Contents
   Step 1. – From the Twitter contents, we first capitalized a list of words that are not capitalized (i.e. "of" "the" "and" "but" "in" "is" "a") if they have a capitalized word on both sides of them.
   Step 2. – Words that are all capitalized (i.e. UPPERCASE) are converted to just capitalized on the starting letter (i.e. Uppercase).

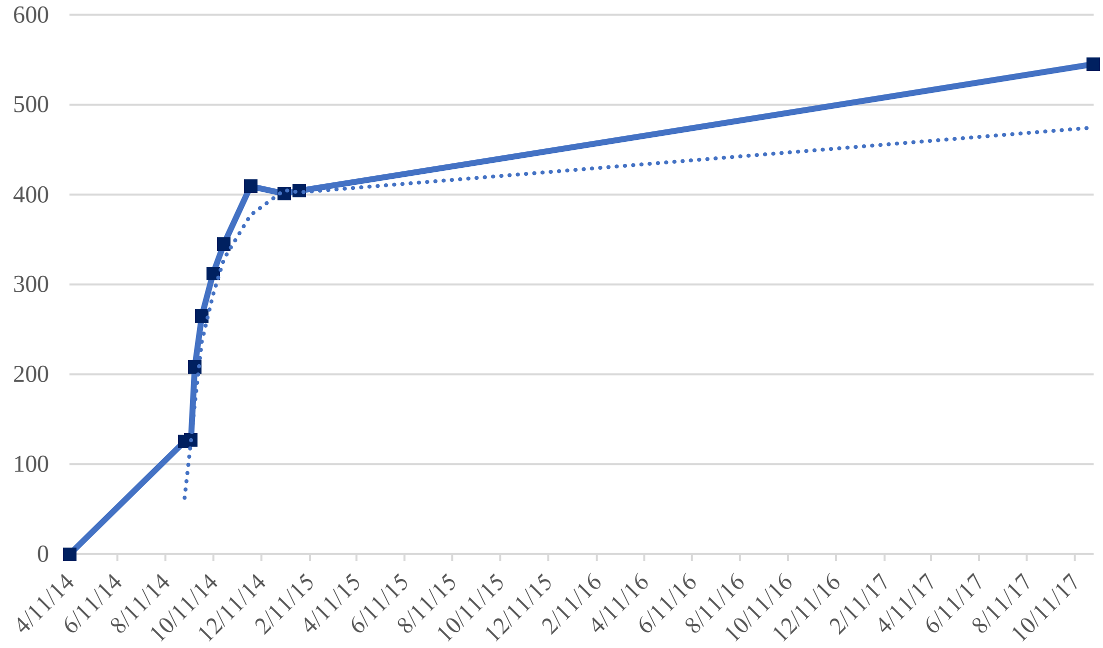
2. Extracting Author Names from the Text
   Step 3. – Author names are then identified as a sequence of Capitalized Words followed by "'s" or preceded by "by", "from".

3. Extracting Book Titles from the Text
   Step 4. – Book titles are those that are between "" or Between '', so "This is a title" and 'This Is A Title'. However, "" which contains "." should be considered as a quotation and dropped (i.e. "This is a title." should be dropped.)
   Step 5. – Book titles are also a sequence of capital words – not just one capitalized word.
   Step 6. – When a sequence of capital words are two capitalized words followed by "'s" preceded by "by" or "from", these are author names, not book titles.
   Step 7. Drop common verbs in the beginning of the tweet that are capitalized (i.e. Follow, Read, etc.). They should not be considered as part of the book title.

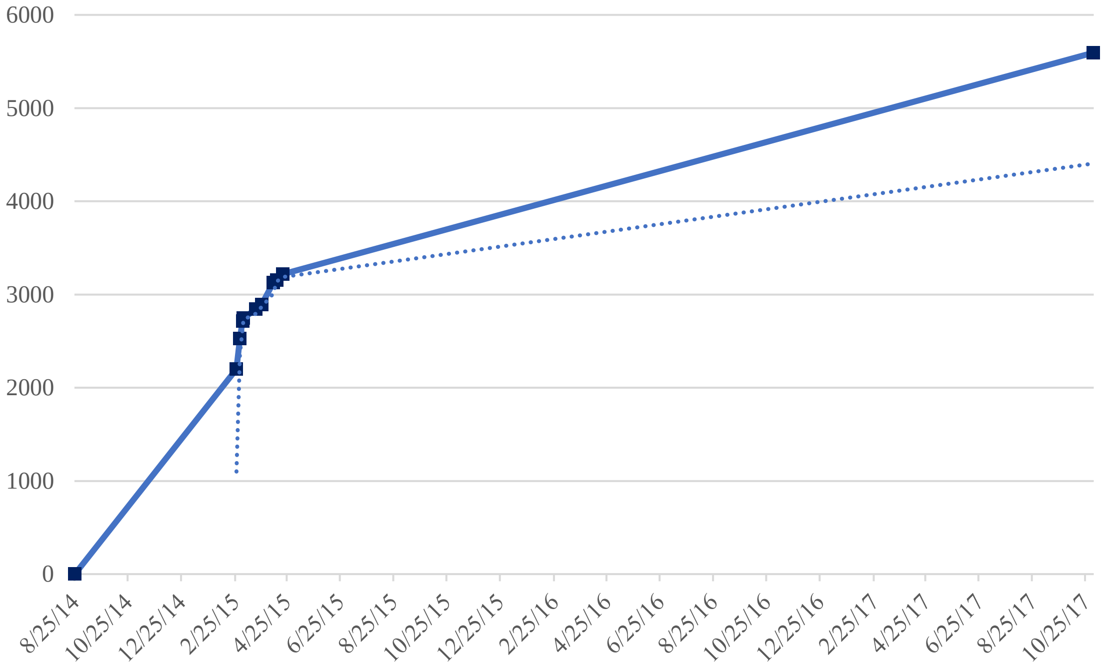
4. Matching Exact Author and Book Titles
   Author names and book titles that are extracted from the Twitter contents in through *Step 3-Step 7* are matched to the author names and book titles of the Amazon eBooks.

5. Fuzzy Matching Exact Author and Book Titles
   Because there are several ways in which a book title is discussed on Twitter (i.e. "Thrid Shift – Pact" by Hugh Howey as "3rd Shift – Pact"), the book titles extracted from the Twitter contents are fuzzy matched to the actual book titles using the Python *Fuzzywuzzy* library.

Appendix B. Number of Followership Changes between the Time of Retweet and the Time of Collection



Number of Followers for @SMoore_Author[8]



Number of Followers for @AtriaIndies

[8] Dotted line shows the moving average

References

Abrahamson, E., & Rosenkopf, L. (1997). Social Network Effects on the Extent of Innovation Diffusion: A Computer Simulation. *Organization Science*, *8*(3), 289–309.

Almandoz, J. (2012). Arriving at the starting line: The impact of community and financial logics on new banking ventures. *Academy of Management Journal*, *55*(6), 1381–1406.

Armstrong, M. (2006). Competition in two-sided markets. *RAND Journal of Economics, 37*(3), 668–691.

Badham, J., & Stocker, R. (2010). The impact of network clustering and assortativity on epidemic behaviour. *Theoretical Population Biology*, *77*(1), 71–75.

Basuroy, S., Chatterjee, S., & Ravid, S. A. (2003). How Critical Are Critical Reviews ? The Box Office Effects of Film Critics, Star Power, and Budgets. *Journal of Marketing*, *67*(October), 103–117.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, Vol. 117, pp. 497–529. US: American Psychological Association.

Benner, M. J., & Waldfogel, J. (2016). The Song Remains the Same? Technological Change and Positioning in the Recorded Music Industry. *Strategy Science*, *1*(3), 129–147.

Berger, J. (2011). Arousal Increases Social Transmission of Information. *Psychological Science*, *22*(7), 891 –893.

Berger, J. (2014). Word of mouth and interpersonal communication : A review and directions for future research. *Journal of Consumer Pyschology*, *4*, 586–607.

Berger, J., & Heath, C. (2007). Where Consumers Diverge from Others : Identity Signaling and Product Domains. *Journal of Consumer Research*, *34*(2), 121–134.

Berger, J., & Milkman, K. L. (2012). *What Makes Online Content Viral ? XLIX*(April), 192–205.

Berger, J., & Schwartz, M. (2011). What Drives Immediate and Ongoing Word of Mouth? *Journal of Marketing Research*, *48*(October), 869–880.

Bernstein, M. (1997). Celebration and Suppression: The Strategic Uses of Identity by the Lesbian and Gay Movement. *American Journal of Sociology*, *103*(3), 531–565.

Bourdieu, P. (1997). The forms of capital. In W. A. M. Halsey A.H., Lauder H., Brown P. (Ed.), *Education, Culture, and Society*. Oxford, UK: Oxford University Press.

Brown, J., Broderick, A. J., & Lee, N. (2007). Word of mouth communication within online communities: Conceptualizing the online social network. *Journal of Interactive Marketing*, *21*(3), 2–20.

Burt, R. S. (1987). Social Contagion and Innovation: Cohesion versus Structural Equivalence. *American Journal of Sociology*, *92*, 1287–1335.

Burt, R. S. (1997). The Contigent Value of Social Capital. *Administrative Science Quarterly*, Vol. 42, pp. 339–365.

Burt, R. S. (1999). The Social Capital of Opinion Leaders. *Annals*, *566*, 37–54.

Buss, D. M. (1990). The Evolution of Anxiety and Social Exclusion. *Journal of Social and Clinical Psychology*, *9*(2), 196.

Carroll, G. R. (1985). Concentration and Specialization : Dynamics of Niche Width in Populations of Organizations. *Merican Journal of Sociology*, *90*(6), 1262–1283.

Carroll, G. R., & Swaminathan, A. (2000). Why the Microbrewery Movement? Organizational Dynamics of Resource Partitioning in the U.S. Brewing Industry. *American Journal of Sociology*, *106*(3), 715–762.

Centola, D. (2010). Supporting Online Material for The Spread of Behavior in an Online Social Network Experiment. *Science*, *329*(5996), 1194–1197.

Centola, D., & Macy, M. (2007). Complex Contagions and the Weakness of Long Ties. *American Journal of Sociology*, *113*(3), 702–734.

Chen, Z., & Berger, J. (2013). When, Why, and How Controversy Causes Conversation. *Journal of Consumer Research*, *40*(3), 580–593.
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.
Childress, C. C., & Friedkin, N. E. (2012). Cultural Reception and Production: The Social Construction of Meaning in Book Clubs. *American Sociological Review*, *77*(1), 45–68.
Cialdini, R. B. (1984). *Influence: Science and Practice*. Needham Heights, MA: Allyn & Bacon.
Coleman, J. (1966). *Equality of educational opportunity*. Washington D. C.: U. S. Government Printing Office.
Coleman, J. (1988). Social Capital in the Creation of Human Capital. *American Journal of Sociology, 94*, 95–120.
Coleman, J. (1990). *Foundations of Social Theory*. Cambridge, MA: Belknap Press.
Coleman, J., Katz, E., & Menzel, H. (1957). The Diffusion of an Innovation Among Physicians. *Sociometry, 20*, 253–270.
Crane, D. (1999). Gender and Hegemony in Fashion Magazines. *The Sociological Quarterly*, *40*(4), 541–563.
Evans, D. S., Hagiu, A., & Schmalensee, R. (2006). *Invisible Engines: How Software Platforms Drive Innovation and Transform Industries*. Cambridge, MA: The MIT Press.
Freeman, J., & Hannan, M. T. (1983). Niche Width and the Dynamics of Organizational Populations. *American Journal of Sociology*, *88*(6), 1116–1145.
Freeman, L. C. (1978). *Centrality in Social Networks Conceptual Clarification*. *1*, 215–239.
Friedkin, N. E. (1993). Structural Bases of Interpersonal Influence in Groups: A Longitudinal Case Study. *American Sociological Review*, *58*(6), 861–872.
Friedkin, Noah E., & Johnsen, E. C. (2013). *Social Influence Network Theory: A Sociological Examination of Group Dynamics*. New York, NY: Cambridge University Press.
Gans, J. S., & Stern, S. (2003). The product market and the market for "ideas": commercialization strategies for technology entrepreneurs. *Research Policy*, *32*(2), 333–350.
Gargiulo, M., Ertug, G., & Galunic, C. (2009). The Two Faces of Control: Network Closure and Individual Performance Amongst Knowledge Workers. *Administrative Science Quarterly*, *54*(2), 299–333.
Gawer, A. (2014). Bridging differing perspectives on technological platforms: Toward an integrative framework. *Research Policy*, *43*(7), 1239–1249.
Gilbert, E., & Karahalios, K. (2009). Predicting Tie Strength with Social Media. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 211–220. New York, NY, USA: ACM.
Golan, G., & Wanta, W. (2001). Second-level agenda setting in the new hampshire primary: A comparison of coverage in three newspapers and public perceptions of candidates. *Journalism and Mass Communication Quaterly*, *78*(2), 247–259.
Golder, S. A., & Macy, M. W. (2014). Digital Footprints: Opportunities and Challenges for Online Social Research. *Annual Review of Sociology*, *40*, 129–152.
Granovetter, M. (1978). Threshold Models of Collective Behavior. *The American Journal of Sociology*, *83*(6), 1420–1443.
Granovetter, M. S. (1973). The Strength of Weak Ties. *American Journal of Sociology, 78*(6), 1360–1380.
Greco, A. N., Milliot, J., & Wharton, R. M. (2014). *The Book Publishing Industry*. New York, NY: Routledge.
Greve, H. R., & Song, S. Y. (2017). Amazon Warrior: How a Platform Can Restructure Industry Power and Ecology. *Advances in Strategic Management (Entrepreneurship, Innovation, and Platforms), 37*, 299–335.
Hannan, M. T. (2005). Ecologies of Organizations : Diversity and Identity. *Journal of Economic*

Perspectives, *19*(1), 51–70.
Hannan, M. T., & Freeman, J. (1977). The Population Ecology of Organizations. *American Journal of Sociology*, *82*, 929–964.
Hedström, P., Sandell, R., & Stern, C. (2000). Mesolevel Networks and the Diffusion of Social Movements: The Case of the Swedish Social Democratic Party. *American Journal of Sociology*, *106*(1), 145–172.
Katz-Gerro, T. (2004). Cultural consumption research: review of methodology, theory, and consequence. *International Review of Sociology*, *14*(1), 11–29.
Katz, E., & Lazarsfeld, P. F. (1955). *Personal influence: the part played by people in the flow of mass communications.* New York, NY, US: Free Press.
Keele, L., & Kelly, N. J. (2006). Dynamic Models for Dynamic Theories : The Ins and Outs of Lagged Dependent Variables. *Political Analysis*, *14*(2), 186–205.
Keeling, M. J., & Eames, K. T. D. (2005). Networks and epidemic models. *Journal of the Royal Society Interface*, *2*(4), 295–307.
Kotha, S., & Basu, S. (2011). Amazon and eBay: Online retailers as market makers. In G. G. Hamilton, B. Senauer, & M. Petrovic (Eds.), *The market makers: How retailers are reshaping the global economy*. Oxford, UK: Oxford University Press.
Kwak, H., Lee, C., Park, H., & Moon, S. (2010). What is Twitter, a Social Network or a News Media? *Proceedings of the 19th International Conference on World Wide Web*, 591–600. New York, NY, USA: ACM.
Leskovec, J., & Horvitz, E. (2008). Planetary-Scale Views on an Instant-Messaging Network. *Proceedings of the 16th International Conference on World Wide Web*, 1–28.
Lounsbury, M., & Glynn, M. A. (2001). Cultural entrepreneurship: Stories, legitimacy, and the acquisition of resources. *Strategic Management Journal*, *22*(6–7), 545–564.
Luminet, O., Manstead, A. S., & Rimé, B. (2000). Social sharing of emotion following exposure to a negatively valenced situation. *Cognition and Emotion*, *14*(5), 661–688.
Marquis, C., & Battilana, J. (2009). Acting globally but thinking locally? The enduring influence of local communities on organizations. *Research in Organizational Behavior*, *29*, 283–302.
Miner, A. S., & Mezias, S. J. (1996). Ugly Duckling No More: Pasts and Presents of Organizational Learning Research. *Organization Science*, *7*(1), 88–99.
Mudambi, S. M., & Schuff, D. (2010). What Makes a Helpful Online Review? A Study of Customer Reviews on Amazon.com. *MIS Quarterly*, *34*(1), 185–200.
Muller, E., & Peres, R. (2019). The effect of social networks structure on innovation performance: A review and directions for research. *International Journal of Research in Marketing*, *36*(1), 3–19.
Newman, M., Barabási, A.-L., & Watts, D. J. (2006). *The Structure and Dynamics of Networks*. Princeton, New Jersey: Princeton University Press.
Oliver, R. L., & Swan, J. E. (1989). Consumer Perceptions of Interpersonal Equity and Satisfaction in Transactions: A Field Survey Approach. *Journal of Marketing*, *53*(2), 21–35.
Onnela, J.-P., Saramäki, J., Hyvönen, J., Szabó, G., Lazer, D., Kaski, K., … Barabási, A.-L. (2007). Structure and tie strengths in mobile communication networks. *Proceedings of the National Academy of Sciences*, *104*(18), 7332–7336.
Pathan, S. (2009). Strong Boards, CEO Power and Bank Risk-Taking. *Journal of Banking and Finance*, *33*(7), 1340–1350.
Podolny, J. M., & Baron, J. N. (1997). Resources and Relationships : Social Networks and Mobility in the Workplace. *American Sociological Review*, *62*(5), 673–693.
Pollock, T. G., & Rindova, V. P. (2003). Media Legitimation Effects in the Market for Initial Public Offerings. *Academy of Management Journal*, *46*(5), 631–642.
Rainie, L., & Wellman, B. (2012). *Networked: The New Social Operating System*. Cambridge, MA: MIT Press.

Rao, H., Monin, P., & Durand, R. (2003). Institutional Change in Toque Ville: Nouvelle Cuisine as an Identity Movement in French Gastronomy. *American Journal of Sociology*, *108*(4), 795–843.

Raridon, A., & Mix, T. L. (2016). “That’s not Grassfed!” Identity Formation, Maintenance, and Co-Optation in Oklahoma’s Grassfed Livestock Movement. *Sociological Inquiry*, *86*(2), 141–165.

Rochet, J.-C., & Tirole, J. (2008). Platform Competition in Two-Sided Markets. *Journal of the European Economic Association*, *1*(4), 990–1029.

Rodan, S., & Galunic, C. (2004). More Than Network Structure: How Knowledge Heterogeneity Influences Managerial Performance and Innovativeness. *Strategic Management Journal*, *25*(6), 541–562.

Romero, D. M., Meeder, B., & Kleinberg, J. (2011). Differences in the Mechanics of Information Diffusion Across Topics: Idioms, Political Hashtags, and Complex Contagion on Twitter. *Proceedings of the 20th International Conference on World Wide Web*, 695–704. New York, NY, USA: ACM.

Salganik, M. J., Dodds, P. S., & Watts, D. J. (2006). Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market. *Science*, *311*(5762), 854–856.

Stiglitz, J. E. (1989). Markets, Market Failures, and Development. *The American Economic Review*, *79*(2), 197–203.

Strang, D., & Soule, S. A. (1998). Diffusion in Organizations and Social Movements: From Hybrid Corn to Poison Pills. *Annual Review of Sociology*, *24*(1), 265–290.

Tarrow, S. (1989). *Democracy and Disorder*. Oxford, UK: Oxford University Press.

Toubiana, M., & Zietsma, C. (2016). The Message is on the Wall? Emotions, Social Media and the Dynamics of Institutional Complexity. *Academy of Management Journal*, *60*(3), 922–953.

Tybout, A. M., Calder, B. J., Sternthal, B., Tybout, A. M., Calder, B. J., & Sternthal, B. (2019). *Using Information Processing Theory to Design Marketing Strategies*. *18*(1), 73–79.

Van den Bulte, C., & Lilien, G. L. (2001). Medical Innovation Revisited: Social Contagion versus Marketing Effort. *American Journal of Sociology*, *106*(5), 1409–1435.

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146–1151.

Warde, A. (1994). Consumption, Identity-Formation and Uncertainty. *Sociology*, *28*(4), 877–898.

Watts, D. J. (1999). Networks, Dynamics, and the Small-World. *American Journal of Sociology*, *105*(2), 493–527.

Watts, D. J. (2002). A Simple Model of Global Cascades on Random Networks. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(9), 5766–5771.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of “small-world” networks. *Nature*, *393*(6684), 440–442.

Weber, K., Heinze, K. L., & DeSoucey, M. (2008). Forage for Thought: Mobilizing Codes in the Movement for Grass-fed Meat and Dairy Products. *Administrative Science Quarterly*, *53*(3), 529–567.

Weick, K. E. (1995). *Sensemaking in Organizations*. Thousand Oaks, CA: Sage.

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT Press.

Wyatt, B. R., & Badger, D. P. (1984). How Reviews Affect Interest in and Evaluation of Films. *Journalism Quarterly*, *61*(4), 874–878.