Growing Influence

Zheng Gong[*]

April 22, 2021

Abstract

How do new firms grow and get established through their product design? I answer this question with theoretical modelling and empirical findings on online influencer market. I collect a new blog posts data from *Wechat Official Account*. Using machine learning methods to categorize topics and identify advertisements in the historical publications of the largest 1002 influencer accounts, I find that entrants to the influencer market start their career by specializing in niche topic with minimal advertisements. Over time they cover broader topics and the extent of sponsored advertisement grows. To understand the underlying mechanism, I develop a dynamic reputation model in which an influencer faces audiences with hetereogeneous tastes, and can choose the fraction of posts allocated to different topics and advertisements over time. The model delivers further implications on how influencers react to exogenous shocks to their own reputation, and to the preferences of audiences, and I find empirical evidence that supports these predictions.

[*]Gong: zheng.gong16@rotman.utoronto.ca. I am especially grateful to Heski Bar-Isaac, Matthew Mitchell and Ignatius Horstmann for their continued discussion, helpful comments and guidance. Moreover, I would like to thank Bernardo Blum, Laurent Cavenaile, Ambarish Chandra, Judith Chevalier, Avi Goldfarb, Aibo Gong, Tanjim Hossain, Guangrui Li, Matthew Osborne, Gabor Virag Xiao Yang, Yanyan Yang and Ruizhi Zhu for their time and helpful comments. I thank Xiaoyi Du for her excellent research assistance and New Rank Co. for their assistance in acquiring data. I thank Rotman China Research Intiative for generous financial support. All errors are my own.

1 Introduction

It is difficult for an entrant to enter a market and become established. One major problem is informational: even if a new entrant is capable of providing good products, it is difficult to persuade consumers to try his products and recognize the quality of the entrant just because he has no reputation in the market. But to gain reputation, an entrant needs to attract consumers in the first place. Therefore, new entrants are in a disadvantaged position when competing with star firms, who have already established reputation in the market.

Economists has long realized this informational problem for new entrants (Bain, 1951), and it is sometimes referred to as a ”cold-start problem” in the literature (e.g., Hui et al. (2020), Aghion et al. (2009)). This problem is more severe in markets where some traditional solutions, like free samples or advertisement campaigns, do not apply. A notable example is the market of online influencers, like YouTubers and bloggers, who are content providers on social media platforms that gain subscribers through the high quality contents that they produce or share. The online influencer market is rapidly growing,[1] but attention is hard to gain, and success is rare in this market. For example, Bärtl (2018) suggests that the earning of 97% YouTubers are below the US poverty line, and 85% of all views go to the top 3% of YouTubers, which has raised concerns among the industry practitioners that such inequality may hinder the entry of new influencers.[2] One aspect that makes it hard for new entrants to compete with star incumbents in the influencer market is that, on most social network platforms, influencers do not charge for contents, and advertisement campaigns are unavailable. The only aspect that an influencer controls is the choice of what content to provide to his audience: that is, the design of the product he offers. This pbservation leads to natural questions: can new entrants alleviate the cold start problem by wisely designing their products? And how does the design of the products evolve as entrants begin to gain reputation and grow over time? On a related note, does the evolution in product design impact how consumers perceive and rate products?

Among various social media, blog platforms are perfect to answer these questions, as blog

[1] The annual growth rate of influencer market in 2016-2020 is estimated to be around 50% (Hub, 2021).

[2] https://www.washingtonpost.com/news/the-switch/wp/2018/03/02/why-almost-no-one-is-making-a-living-on-youtube/

posts are mainly in text and there are mature machine learning techniques to extract information from text data. In this paper, I collect the historical publication data on *Wechat Official Account platform*, the largest blog platform in China, and I apply natural language processing techniques to classify blog contents, and there by, to quantitatively analyze the evolution of contents over time. I find that an influencer's contents become more dispersed in topics and contain more advertisements as he grows. To understand the underlying mechanism, I propose a theoretical framework where an influencer designs the allocation of his contents into different topics and sponsored advertisements over time. The model derives further implications on the reactions of influencers to exogenous shocks on reputation and preference of audiences, and it allows us to understand dynamics in the ratings of contents. I provide empirical evidence that supports these implications.

In the data, I download the historical publications of 1002 most influential official accounts on Wechat Official Account platform, with 5.53 million observations and the time spans from September 2014 to December 2019. I observe the contents of blog posts, the publish time, and likes and clicks counts for each article. In order to consider product design among blog posts, I use natural language processing techniques to analyze the posts data. I employ Latent Dirichlet Allocation (LDA) method as the topic model, which maps high dimensional text data - each blog post - into a vector of scores over different topics. And I use both Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) models to identify advertising contents in the data. Equipped with the quantification of content design, a natural question to ask is: whether bloggers at different stage of career systematically choose different product design and why? I find that the content design evolves over time in two aspects. First, an influencer covers more dispersed topics in his contents when he gets established. I measure the dispersion of blog posts in the topics space at influencer - age level. I find that the topic dispersion increases by 8.7% annually for an average influencer. Second, influencers embed more advertisements in their contents as they grow. Again, I measure the share of advertising posts at influencer - age level, and I find that the ads share increases by 46% annually.

What drives such dynamics in content design? I propose a theoretical model to understand the mechanism. An influencer faces two types of horizontally differentiated audiences over time, and each audience has an idiosyncratic match value with the influencer. At each period of time, the influencer chooses the horizontal design of his content, which varies from the most *niche* (which favors one type of the audiences more than the the other) to the most *broad* (which treats both types of audiences equally). The influencer also chooses the extent of adverting embedded in the content, which benefits the influencer and harms the audiences. The content reaches all of his old subscribers and some new audiences at each period of time, and audiences make content consumption decisions. The influencer has an underlying type of being a good content provider or not, which is initially unknown to all the players, and over time his type is gradually revealed to the public through the quality realization of his past contents.

The first prediction from the model is that influencers tend to start their career by offering more niche contents, and over time they gradually broaden the content. The intuition behind it is that the influencer faces competition from the market, which is reflected by the audiences' outside options. When an influencer has not built reputation yet, it is relatively difficult to convince audiences to consume his content rather than their outside options, which might be an established influencer's contents. By offering niche content, the influencer makes the distribution of audiences' expected utility more dispersed. That is, making some audiences more likely love him and others more likely hate him. Those who really love him consume his contents even though he has not established reputation yet, and the influencer can therefore start to accumulate reputation and subscribers. While young influencers gain from the up-side risk, the influencers that have already established their reputation would avoid losses from the down-side risk by reducing the dispersion of audiences' utility. Therefore, over time an established influencer would design more broad contents, so to encourage more types of subscribers to consume his content and reach more broad new audiences. This finding not only consist to my data, but also echos the success stories of many real world influencers. For example, Marques Brownlee, a technology YouTuber, describes the evolution of his contents: "(I) snowballed into just making all kinds of videos with the laptop, and then the software the mouse... and my channel just turned into a tech YouTube channel... And we're launching one other channel with a lot more casual content."[3] Another YouTuber,

[3]https://www.theverge.com/22231657/mkbhd-marques-brownlee-interview-youtube-creator-influencer-decoder

Chris Ballinger, states: ”My channel started out as a place for me to post magic videos. Eventually we started posting one magic video a week and one random video to try out new things.”[4] When comparing ”single game streaming” with ”variety streaming”, Stream Edge, a Twitch blog suggests: ”One of the pros for single game playing is that it’s easier to build audiences with one game... And one of the cons of variety streaming is that it takes a lot more time to grow at the beginning.”[5]

The second prediction from the model is that while the influencers gain reputation and subscribers, the share of advertisements in their contents also grow over time. There are two reasons behind this. First, entrants have low initial reputation, and the only way to attract audiences is to minimize the advertisements. As an influencer accumulates reputation for providing good contents, audiences have higher expectation for the quality of his contents, and can bear more advertisements. Second, influencers accumulate subscribers over time. With more subscribers, an influencer has higher incentive to provide more advertisements and monetize from the existing subscribers, rather than a lower level of advertisements and attract more new subscribers. This strategy is also consistent with observations of the market. As John Koetsier comments on Forbes, ”for smaller YouTube creators, being ad-free can be a competitive advantage as they climb the long hill to having a monetizable YouTube channel with a reasonable amount of revenue.”[6]

The model provides us a framework to understand the impacts of some exogenous shocks to the environment, and find additional supporting evidence of the central economic force of the model. First, I examine the impact of a shock to the reputation of influencers in this market, the original tag function. Wechat invites official accounts that they see as ”high quality” to access to the original tag function, which allows influencers to label their posts as ”original”. Getting this invitation signals the high quality of an influencer to the audiences, and boosts his reputation. I apply a coarsened exact matching (CEM) approach to examine the impact of original tag invitation, and, as suggested by the theory, I find that original tag brings a significant and sizeable

[4] https://mediakix.com/blog/chris-ballinger-youtuber-interview-family-vlogger/

[5] https://medium.com/@streamedgeanalytics/the-pros-and-cons-of-single-game-vs-variety-streaming-844eb9b5c73d

[6] https://www.forbes.com/sites/johnkoetsier/2020/11/18/youtube-will-now-show-ads-on-all-videos-even-if-creators-dont-want-them/?sh=27368b3e4913

increment in topics dispersion and advertisements in blog posts. The impact is equivalent to an additional 7.5-9.3 months of age in topics dispersion and 14 months of advertising increment for an average influencer.

Second, I examine the how influencers at different stages react to shocks in audiences' preference. Sometimes big events emerge in society, like Covid-19 or the Capitol riot, and have widespread impact. These events can be interpreted in the model as exogenous shocks that change preferences, as all the audiences would like to know something about these new events to some extent. Should a young influencer divert a lot of his routine contents into those events to try to attract attention and gain subscribers? Naively this may be a good strategy to attract eyeballs. But following the logic of the model, I show that the answer is no and, again, find supporting evidence in the data. I extend the model by assuming that an influencer can divert some fraction of his contents to a new general-interest event, and audiences of both types gain from it. A greater fraction of content on the new event benefits both types readers, but reduces audiences' gain from the routine contents. I show that compared to young influencers, old influencers respond more to the new event and allocate a larger share of content to it. The idea is that an old influencer designs broader contents and aims to grasp attention from both types of audiences. The content on this new event interests both types of his audiences and therefore, in any case, aligns with his previous content design incentive. On the other hand, a young influencer tends to choose a more niche content and targets to a narrow range taste of audiences. This creates misalignment of the new event and his previous niche content design.

Empirically, I take advantage of the Meng Wanzhou arrest event, which happened on Dec 1st, 2018, as a natural experiment. As Huawei, the company where Meng Wanzhou served as CFO is the biggest tech company in China, this case and its impacts are widely discussed on the Chinese social media at that period of time. Besides, Meng Wanzhou case is unexpected to everyone, and it is easy and clear to label Meng Wanzhou related blog posts. All above make Meng Wanzhou event the best candidate for a preference shock to audiences. I compare the official accounts that are younger than 1 year at that moment and the older accounts, and apply a diff-in-diff approach to compare the changes in fractions of Meng Wanzhou related articles post the shock. The data shows that over a 30-days post trend, the old influencers publish 32% more Meng Wanzhou related articles in fraction than the young ones.

Last, rating of contents is an important signal of quality, but my model suggests that ratings should be also endogenously affected by the design of the contents and the composition of audiences. Are ratings biased and in what direction? How does the distribution of ratings evolve over an influencer's life cycle and why? To answer these questions, I extend the baseline model by assuming that better content gives readers a higher perceived value, and is shared more to subscribers' friends, thereby attracting more non-subscriber audiences. Two forces both reduce the average rating over time. First, since advertisements grows, the vertical quality decreases over time and audiences' perceived valuations decrease for a given audience population. Second, an influencer's subscribers are not randomly drawn from the general population, but are self-selected among those who have a high match value with the influencer. Since a young influencer specializes in a niche topic and has low reputation, the selection bias for a young influencer is greater than that of an old influencer, which leads to more upward biases in the average rating. Interestingly, due to the selection bias, I find that ratings are more upward-biased for low quality contents, especially for the young influencers. This implies that for a given influencer, the distribution of ratings become more dispersed over time, and ratings gradually become better signals of the true quality of contents. Empirically, I use the likes-clicks ratio of each article as a proxy of rating to examine its dynamics. I find that average ratings significantly drop by 7.2% per year over the initial two years of an influencer's life cycle. Moreover, I compare the descending rate of ratings of different quantiles of blog posts in clicks, for each influencer over time. I find that the decreasing in rating is mostly explained by the low clicks contents. the decreasing rate of ratings for top quarter quantile articles is only 1.3% per year, compared to 13% for the bottom quarter quantile articles.

The findings in my model and data are relevant not only for Wechat or the influencer industry, but may also shed light on other markets, especially those industries with horizontally differentiated consumers where product design is relatively flexible (e.g., PC game producers, musicians, game streamers, etc.). When a new enterpreneur enters those industries, he might naively design a mainstream product that is all things to all people. Our model suggests, however, that this may not be a good choice. Even if he is a *de facto* high ability enterpreneur and produces good products, since he is new and has no reputation for that, all consumers may hesitate to discover his mainstream product. On the other hand, designing a niche product helps, as it attracts zealot fans whose tastes match exactly with the niche design, and the new entrant can begin to accumulate reputation and loyal consumers.

The rest of the paper unfolds as follows. In Section 2 I summarize the related literature. Section 3 describes the institutional background and the data I use. Section 4 presents some preliminary findings from the data. Section 5 presents the model. In Section 6 I derive some further implications from the model, and provide the empirical evidence. Last I conclude in Section 7. Appendix A, B and C contain proofs, Figures and Tables, and numerical examples respectively.

2 Related Literature

This research is most directly related to the abundant literature that studies the influencer contents. There are two main streams of literature in the influencer content design and my paper relates to both sides. The first stream aims to understand an influencer's trade-off in providing genuine product recommendations versus sponsored advertisements, and the policy implications of mandatory disclosure regulations. Mitchell (2020) studies a dynamic game between an influencer and an audience. The influencer trades-off between good advice to gain reputation as a commitment type, or advertisements to gain revenue over time. Pei and Mayzlin (2019) and Jansen and Williams (2021) study a similar problem statically but internalize the relationship between the influencer and the sponsor of advertisement. Ershov and Mitchell (2020) empirically tests the treatment effect of Ads disclosure regulation on the amount of advertisements. Fainmesser and Galeotti (2018) examines a competitive market where influencers are vertically differentiated and mix advertisements with authentic contents.

The second line of researches studies the choice of non-advertising contents by influencers and other media. Sun and Zhu (2013) examine the effect of platform sharing advertising revenues with the influencers. They find that influencers who receive the treatment publishes more contents in

”popular” topics, with higher quality. Seamans and Zhu (2014) show that the entry of Craigslist increases content differentiation between newspapers. Kerkhof (2020) shows that an increase in advertisements on YouTube drives YouTube contents to be more diverse. Chen and Suen (2019) build a theoretical model to study the accuracy of news in a competitive environment. Katona et al. (2017) theoretically studies the competition of news providers on different topics. They find that branded media who have more loyal audiences choose ”safe topics” of news, while unbranded media choose more risky topics. My paper is different from most of the past research on influencer contents in the sense that I focus on the life cycle of influencers, and how the content design evolves over an influencer's life cycle. This stream of research is typically static or examines the treatment effects of some policies.

This paper is also closely related to the firm and career life cycle literature. A broad literature, date back to Shapiro (1983), focuses on the informational problem of young firms. That is, high-quality young sellers cannot be distinguished from low quality ones because they have not established reputation yet, which is also the problem that young influencers face in this paper. Literature on this question typically propose the design of market and reputation system as a solution. For example, Klein et al. (2016) studies the effect of a reputation system policy on Ebay that aims to improve quality transparency. Barach et al. (2020) conduct an experiment to show that platforms can actively steer consumers to buy from new sellers by recommendation and guarantees. Bai et al. (2020) conduct experiment to show that initial demand shock alleviates the cold-start problem. Hui et al. (2020) show that a less history-dependent reputation system helps the young sellers. Li et al. (2016) examines the signaling effect of a reputation system that allows sellers to buy feedback. I add to this literature by showing that not only the platform, but the sellers can also actively overcome the cold start problem by choosing a niche product design. Another stream of literature focuses on the ”exploration versus exploitation” mechanism, such that agents (either firms or workers) explore more options in the early stage of their life cycle, and exploit the best alternative (for example, Jovanovic, 1979, Sicherman and Galor, 1990 and March, 1991). Interestingly, my model suggests the opposite such that influencers specialize initially, rather than later. My data appear to support my theory, which suggests that ”exploration and exploitation” may not be the first order driving force of the content topics choice.

On the theoretical side, the intuition of my model is aligned with the horizontal product design literature following seminal papers like Lewis and Sappington (1994) and Johnson and Myatt (2006). In this line of research, the product designer can choose product design from a series of demand rotations, varying from the most niche to the most broad. Johnson and Myatt (2006) show that under mild conditions, the optimal product design is either most niche or most general. In Bar-Isaac et al. (2012), the authors find that high quality firms in the search market choose the most broad design and low quality firms adopts the most niche design. Anderson and Renault (2009) show that weak firms tend to apply the information strategy that spreads the match value distribution from the consumers. My paper has similar result that a weak influencer (who has low reputation) adopts most niche design. However, since influencers accumulate past subscribers, the content design may not directly jump from the most niche to the most broad,[7] but gradually broadening over time.

Last, my paper is related to the literature that examines the product ratings and the dynamics of ratings. Bondi (2019) is the closest paper such that in both papers, niche product leads to a stronger self-selection effect and biased high ratings. My model is different in the sense that content design in my model is endogenous, and Bondi (2019) focuses on the consumer side and treats product design as exogenously given. Vaccari et al. (2018) has a similar result that quality differences are underestimated by ratings, though through a different mechanism, whereby individual preferences are reference dependent. On the dynamics of rating, Filippas et al. (2019) shows the existence of ”rating inflation” on platforms. My paper, however, shows both theoretically and empirically that for each influencer, average ratings drop over time. This implies that Filippas et al. (2019) may even underestimate the actual extent of rating inflation.

[7]The bang-bang product design (either most niche or most broad) is a common feature in this literature with rare exceptions like Bar-Isaac et al. (2021), who highlight that better firms choosing broader design does not rely on firms choosing only extremal design.

3 Institutional Background and Data

3.1 Background

The empirical setting for this paper is *Wechat Official Account platform*, the biggest blog platform in China. The blog market is well-suited to study content design. Contents on the blog market are mostly text, and we have mature machine learning techniques to analyze text data, compared to videos or graphs. Blogs display large horizontal and vertical differentiation, and audiences have a wide range of horizontal tastes in this market. Also unlike social networks like Facebook or Instagram where influencers are the same as any other user with more friends, it is straightforward to distinguish influencers from audiences in the blog market.

The Wechat Official Account platform is hosted on Wechat, which is the leading social media app in China with more than 1.21 billion active users as of 2019.[8] Wechat offers multiple services, including instant messaging, photo and video sharing, financial services, gaming, news and blogging. First launched in 2012, the Wechat Official Account Platform enables its users to setup blogs (called ”official accounts” on WOA platform) on Wechat. Soon it became the most popular blog-like platform in China, with over 20 million official accounts, and all the Wechat users are also accessed to contents on the WOA platform by default. The articles on this platform generate more than 3 billion views per day, and many popular official accounts have millions of subscribers.

Wechat official Account platform works similar to the news feed on Facebook. Ordinary users can subscribe to the official accounts that they are interested in. Once an official account publishes new articles, its subscribers get notification in the Wechat app. If an influencer publishes more than 3 articles at one time, its subscribers see the titles of the first 3 articles directly, and have to click ”more article(s)” to unfold the titles of the rest of the articles. Subscribers can ”like” the articles that they have read, and they can also share the articles to group chats, private chats, or the ”moment” of Wechat, which functions similar to Facebook timeline such that the sharing is displayed to the user’s friends. See Figure I for a typical interface of Wechat Official Accounts. Unlike some other social media platforms which use algorithmic recommendation system to select news feed, the

[8] https://www.messengerpeople.com/global-messenger-usage-statistics/

displays of official accounts notification and Wechat moment are totally chronological during my data time period. Wechat restricts the official accounts to publish articles only once per day (with rare exceptions for government and media related accounts). But an account can publish at most 8 articles at the same time.

Wechat does not pay official accounts directly. There are three major channels that an official account may profit from. First, official accounts are allowed to publish sponsored advertising posts and get paid from the third party sellers. Wechat does not take shares from those advertising posts. Second, Wechat has its own advertising system. Influencers with more than 500 subscribers can enable advertising banners showing up in the top/middle/bottom of their blog posts, and the advertising revenue is shared between Wechat and the influencer. Third, starting from June 2015, readers are allowed to tip bloggers directly. Yet WOA practitioners claim that the income from tips is much smaller than that from advertisements.[9] In the model, we will focus on the trade off between the advertising contents and normal contents.

Wechat Official Account platform allows a wide range of media in the blog posts, including text, graphs, audios, videos and links to external websites. Yet text is still the dominating media among the blog posts on the platform, which allows us to apply natural language processing techniques to analyze the data.[10] Since Wechat is widely used by ordinary Chinese people, its users have large horizontal differentiation in their taste to contents, and the bloggers on the WOA platform also cover a wide range of topics (see Table II for representative topics and keywords in my data). Wechat allows official accounts to publish original blog posts that they write on their own, or alternatively repost articles from other bloggers. There exists a change in the repost policy. Prior to June 27, 2018, official accounts need to obtain consent from the original blogger to repost an article. After that, all articles are by default allowed to be repost. The name and link to the original blogger will be shown in the repost articles.

[9] https://zhuanlan.zhihu.com/p/37689260

[10] For example, only 2.3% blog posts have less than 50 words in my data.

3.2 Data

The data contains the historical blog posts of 1002 most influential official accounts as of December 2019,[11] with the exception of real-world entity related official accounts (accounts of the government, media, firms, restaurants, etc.). I also exclude those accounts that mainly publish graphs, audios or videos because my empirical analysis relies on NLP techniques. The time span of the data is from Sep 1, 2014 to Dec 10, 2019. For those official accounts that are established after Sep 1, 2014, I observe all of their historical publications. For the rest, I observe only publications after Sep 1, 2014. Table I shows the summary statistics of the data. For each article, I observe all the information of each blog post from a normal reader's view, such as the title, the author (if specified), the full content, the time of publishing, whether it is flagged as original article, and its order in the daily notification. Besides, I also obtain the total number of clicks and likes for each article.

3.3 Processing Text Data

The major part of the Wechat Official Accounts data is the texts of blog posts. But it is difficult to directly utilize the text data in the empirical analysis. To derive quantitative features of the content design dynamics, I employ NLP techniques to analyze two aspects of the data: the topic(s) of each article, and whether an article is an advertising post.

3.3.1 Topic Modeling

To categorize blog posts into topics, I apply an unsupervised machine learning method, latent Dirichlet allocation (LDA; Blei et al., 2003). There are apparent benefits of using the LDA method. LDA is the most commonly used machine learning topic model in text analysis tasks, especially in the marketing and management researches (Blei, 2012, Hannigan et al., 2019, Reisenbichler and Reutterer, 2019). It is highly efficient for big data and data with sparse matrix, like the text data of blog posts. And most importantly, LDA is an unsupervised method, which involves minimum human intervention.

[11] According to the ranking released by the New Rank Ltd. See
https://www.newrank.cn/public/info/list.html?period=month&type=data

Several pre-processing steps are done before I apply the LDA method. First, most of my contents are natural language text data in Chinese. It has no segmentation symbol (like spaces in English) between words, so that I have to segment the Chinese characters into words. For example, the Chinese contents look like "istudyeconomicsintoronto" and to apply LDA, it should be segmented to "i/study/economics/in/toronto". I use Jieba segmentation[12] for this step. Second, the text should be cleaned to remove words that are not informative on topics. I used the standard stop word dictionary that is provided in Jieba project to remove all the stop words (e.g., the, and, at, which) and the punctuations (e.g., ",", "?", "!").

LDA methods requires an input of number of dimensions. In my main analysis, I use the results of 30 topics as this is the largest number that the high weights key words in different topics still display high heterogeneity. For robustness, I also show results with other numbers of dimensions. The LDA method assigns each article k into topic scores over the N topics, $S_k = (s_1, s_2, \ldots, s_N)$, which reflects the intensity of the topics on each article. The normalized score ($\Gamma_k = \frac{S_k}{\sum S_k}$) represents the probabilistic distribution of the article on each topic. Figure II shows the distribution of the highest probability topic of all the articles in my data, and Table II presents the high weight key words for all the 30 topics. The top 3 most frequent topics are clustered in love and marriage, parenting, and entertainment.

3.3.2 Categorizing Advertisements

To label advertising posts, I apply the convolutional neural network (CNN) method to build the classifier of advertisements. The CNN method is widely used in the literature as text classifier, and has been proven to be reliable and efficient by the literature (Georgakopoulos et al., 2018, Wang et al., 2019, Amjad et al., 2019).

CNN method is a supervised machine learning method, which requires labelled contents as the training data to train the model. To limit my own intervention to this step, I outsourced the data label task to a crowd sourcing company in China, Mayi Zhongbao.[13] I hired this company to label 13,500 random articles during July to August 2020. The crowd-sourced labelers are asked if

[12] An open-sourced Chinese words segmentation project. See https://github.com/fxsjy/jieba

[13] See https://www.antzb.com/

the main body of the article is advertisement, and how certain (binary, certain or not so certain) they are for the judgement.

I then use the open-sourced neural network text classifier by Tencent to train the model and predict the rest of the data.[14] I divide the labelled data into 12,000 training set and 1,500 test set. The precision of the test set reaches 92.1% with the CNN method. 11.4% of articles in the whole data are identified as advertisements.

4 Basic Empirical Findings

In this section, I present some empirical findings on the dynamics of content design from the data, which motivate us to build a theoretical framework to understand the underlying mechanism and deliver new implications. More specifically, I look into the dynamics of the following two aspects. First, how disperse is the topics covering in the contents? Second, how much advertisement influencers embed in their contents?

4.1 Dynamics in the Dispersion of Topics

I first consider the dispersion of topics in contents. This is an interesting question to understand, as it reflects the extent of targeted audiences for an influencer. Contents that concentrate in a specific topic interests only a certain group of audiences who favor that topic. If contents cover more dispersed topics, it attracts a wider range of audiences, but everyone may dislike some part of the contents that are not of her taste.

To measure the dispersion of topics in the publication, I assume that the topics that are predicted by the LDA model forms an N-dimension Euclidean space ($N = 30$ in my main result). For each article, the LDA model predicts its scores over N topics, and I map the vector of scores into the N-dimension space as the article coordination in the topic space. For each influencer i, at time t (in month, starting from the first day of my data), I denote the collection of articles as $\mathcal{A}_{it}$. I use the average Euclidean distance of topic scores for all the articles in $\mathcal{A}_{it}$ to the mean to measure the

[14] See https://github.com/Tencent/NeuralNLP-NeuralClassifier. It is an open-sourced text classifier project that is developed by Tencent.

dispersion of topics for influencer i at time t. That is:

$$Dispersion_{it} = \frac{1}{N_{it}} \sum_{k \in \mathcal{A}_{it}} \left| S_k - \frac{1}{N_{it}} \sum_{k \in \mathcal{A}_{it}} S_k \right|$$

where N_{it} is the total number of articles for influencer i at time t. Figure III presents the time trend of topic dispersion that is averaged over the accounts at the same age (month). For robustness, I also use the standard deviation of topic scores for all publications in $\mathcal{A}_{it}$ as an alternative measure of dispersion. The baseline empirical specification to estimate the time trend of dispersion is:

$$\ln(Dispersion_{it}) = \alpha_i + \alpha_k + \beta_1 Age_{it} + \beta_2 InitialDate_i + \epsilon_{it}$$

where α_i is the fixed effect for influencer i, α_k is the fixed effect for topic k, which is the major topic in $\mathcal{A}_{it}$. $InitialDate_i$ controls for the time that each influencer enters the market, and Age_{it} denotes the age of influencer i at time t. Table III presents the results. Column (1) is our main result that measures the topic dispersion by the average Euclidean norm to the center of $\mathcal{A}_{it}$. Column (2) - (7) show that the result is robust. In column (2), I instead use the standard deviation of articles in $\mathcal{A}_{it}$ to measure dispersion. In column (3), I also control for the year fixed effect. In column (4), I control for both year fixed effect and a policy change that took place in Jun. 27, 2018 that makes repost easier for the influencers. In column (5) I exclude influencers that pre-exist in the beginning of the data (Sep 1, 2014),[15] as well as accounts that are younger than 3 months. In column (6) and (7), I use the LDA result of 35 topics and 25 topics respectively. To interpret the economic magnitude, the topic dispersion for an average influencer significantly increases by 0.7% per month, or 8.7% per year.

4.2 Dynamics in the Advertising Contents

Next, I examine the extent of advertising contents over an influencer's life cycle. This is an important decision for influencers, especially bloggers on the Wechat Official Account platform, as

[15] I identify all the accounts that have publication during Sep 1, 2014 - Sep 7, 2014 as pre-exist accounts. There are 96 in total.

advertising is the main source of income for most influencers.

My main identification employs the CNN method as described in Section 3.3.2. As robustness checks, I also apply recurrent neural network (RNN) method and use key words ("purchase", "snap up", "Taobao", etc.)[16] to identify the articles of advertisement. For each influencer i at time t, I calculate the shares of articles that are predicted to be advertisements among all the articles in $\mathcal{A}_{it}$. Figure IV presents the time trend of advertisement shares that is averaged over the accounts at the same age (month). The empirical specification is as follows.

$$ln(AdvShare_{it}) = \alpha_i + \alpha_k + \beta_1 Age_{it} + \beta_2 InitialDate_i \epsilon_{it}$$

Again, α_i is the influencer i fixed effect, and α_k is the fixed effect of the major topic in $\mathcal{A}_{it}$. Table IV shows the regression result of this task. I apply different methods to identify advertisements in column (1) - (3) (CNN, RNN and Keywords), and in column (4) I exclude influencers that pre-exist in the beginning of the data (Sep 1, 2014) and influencers that are too young (less than 3 months). Both CNN and RNN method gives similar prediction, that the fraction of advertisements grows at a rate of about 3.23% per month, or 46% per year.

5 Theoretical Framework

5.1 Benchmark Model Setting

I consider one influencer (he) who faces two types of audiences (she), with type $j = \{A, B\}$. The market operates over an infinite discrete time horizon, $t = 0, 1, 2, 3 \dots$. At each period of time, mass 1 of each type of audience arrives at the market and a fraction of $\gamma \in [0, 1]$ old audiences depreciates. For model tractability, in the benchmark model I assume all agents are myopic.[17] At each period t, the influencer designs the content $\pi_t = (x_t, a_t)$. $x_t \in [0, 1]$ denotes the horizontal design of his content, where $x_t = 0$ favors the type A audiences most and $x_t = 1$ favors the type B

[16] I search in the data with Chinese characters. An article is labelled as advertisement if it contains at least one of the following words: "淘宝", "taobao", "天猫", "手淘", "购买", "抢购".

[17] I provide a numerical simulation of the model with forward looking influencer in the Appendix C. I find that similar to myopic influencers, forward looking influencers also apply a niche to broad content design strategy, and the extent of advertising monotonically increases over time.

audiences the most. We hereafter call $x_t = \{0, 1\}$ the most *niche* design and $x_t = \frac{1}{2}$ the most *broad* design. $a_t \in [0, \infty)$ is the amount of advertisements that is associated with the content, which impacts the vertical quality of the content. A high a_t reduces the utility that both types' audiences gain from consuming the content. The influencer bears an opportunity cost c to produce contents at each period of time and he could decide to quit at any period of time.

The influencer has two underlying types, $k \in \{H, L\}$. All agents do not know the actual type of the influencer, but hold $\theta_0 \in [0, 1]$ belief on the influencer being H type. At each period of time, the influencer produces either Good or Bad content realization. The probability of type θ influencer produces Good content is denoted by λ_k, and we assume $0 \leq \lambda_L < \lambda_H = 1$. That is, a *Bad* content is a perfect bad signal. Once it realizes, the public belief on the influencer's type degenerates to 0.

Audience i's utility from consuming the content is determined by 4 factors: the realization of content quality: {Good, Bad}; the audience's idiosyncratic match value to the influencer u_i; the horizontal location of content, x_t; and the amount of advertisement, a_t. We assume that $u_i \sim \Phi(\cdot)$, where the p.d.f $\phi(\cdot)$ is log-concave and symmetric around the expectation $\mu > 0$.[18] That is, $\phi(\mu - x) = \phi(\mu + x), \forall x$. The payoff for audience i of type A is:

$$W_t^A = u_i - x_t - a_t - \mathbb{1}\{\text{Bad}\}\Delta$$

and $W_t^B = u_i - (1 - x_t) - a_t - \mathbb{1}\{\text{Bad}\}\Delta$, where Δ is the utility difference between a good content and a bad content. Lastly, I assume that Δ and c are large enough, and λ_L is low enough such that the once the influencer is revealed to be of type B, he cannot cover the opportunity cost c and quits the influencer market. This assumption is consistent to recent evidence that the drop out rate in the influnencer market is very high (Bärtl, 2018). The timing of the game at each period t is as follows.

- New audiences enter the market and realize u_i. Everyone update the public belief from last period payoffs (and θ_0 if at $t = 0$).

[18] Many commonly seen distribution functions, like uniform, normal and Laplace distribution satisfies these assumptions. Past literature also have made similar assumptions, like Anderson and Renault (2009).

- The influencer designs content (x_t, a_t) and announces.
- Both the new audiences and the old subscribers decide whether to consume the content or not. If a new audience consumes the content, she becomes a subscriber and stays in the market in the future. Otherwise she quits the market.
- Quality of content is realized and the payoffs are realized for all agents.
- γ fraction of existing subscribers depreciates and quits the market.
- The game proceeds to the next period.

5.2 Benchmark: No Past Subscribers

To illustrate the intuition of this model and highlight the role of the subscriber base, we first study a simple situation where the depreciation rate of subscribers $\gamma = 1$. That is, all the old subscribers depreciate at the end of every period, so that the influencer faces the same demand, mass 1 of each type audiences, at every period of time. The only evolving parameter is the public belief on the influencer being H type, θ_t. Denote ξ_t as the probability that the influencer produces a bad content at this period. We have $\xi_t = (1-\theta_t)(1-\lambda_L)$. Since there exists no past subscriber and two types of audiences are symmetric, w.l.o.g we let $x_t \in [0, \frac{1}{2}]$. The optimization problem for the influencer is:

$$\max_{x_t, a_t} \left[2 - \Phi\left(a_t + x_t + \xi_t \Delta\right) - \Phi\left(a_t + 1 - x_t + \xi_t \Delta\right)\right] a_t \tag{1}$$

and the following Lemma shows that the optimal horizontal design is extreme.

Lemma 1. *For any given advertising level a_t and other parameters, the profit is quasi-convex in the horizontal content design x_t. This leads the influencer to choose extreme design: $x_t = 0$ (most niche) or $x_t = \frac{1}{2}$ (most broad).*

The proof of this result and all the other theory results are in the Appendix A. The intuition here is aligned with Lewis and Sappington (1994): dispersion in demand can be either good or bad to the influencer. When only the audiences who really love the influencer would consume his content, the influencer gains from a larger dispersion, that is, making audiences either really love him or hate him rather than pooling in a tepid attitude. He gains from changing a tepid audience to his die-hard fan, but he does not lose from making a tepid person hate his content, as that person would not consume his content anyway. But if as long as audiences do not hate the influencer, they will consume his content, the influencer would avoid making people hate him. That is, a reduction in the dispersion of demand benefits the influencer.

More specifically, when $a_t + \xi_t \Delta + \frac{1}{2} \le \mu$, it is optimal to set the broad design such that $x_t = \frac{1}{2}$. Otherwise it is optimal to set $x_t = 0$. Note that this depends on the level of a_t. Let $\hat{a}_t = \max\{0, \mu - \frac{1}{2} - \xi_t \Delta\}$, which is the highest a_t such to make the influencer chooses the broad design.

If the influencer chooses a small advertising level and broad product design, the optimal amount of advertisements (without the constraint that $a_t \le \hat{a}_t$) is determined by:

$$\tilde{a}_t(B) = \frac{1 - \Phi\left(\tilde{a}_t(B) + \xi_t \Delta + \frac{1}{2}\right)}{\phi\left(\tilde{a}_t(B) + \xi_t \Delta + \frac{1}{2}\right)} \tag{2}$$

and the amount of advertisements under broad mode is $a_t^*(B) = \min\{\tilde{a}_t(B), \hat{a}_t\}$ Similarly, if the influencer chooses a niche product design, the optimal amount of advertisements (without the constraint that $a_t \ge \hat{a}_t$) is determined by:

$$\tilde{a}_t(N) = \frac{2 - \Phi\left(\tilde{a}_t(N) + \xi_t \Delta + 1\right) - \Phi\left(\tilde{a}_t(N) + \xi_t \Delta\right)}{\phi\left(\tilde{a}_t(N) + \xi_t \Delta + 1\right) + \phi\left(\tilde{a}_t(N) + \xi_t \Delta\right)} \tag{3}$$

and the advertising level under niche mode is $a_t^*(N) = \max\{\tilde{a}_t(N), \hat{a}_t\}$. The influencer's payoff at time t, Π_t^* is determined by:

$$\Pi_t^* = \max\left\{2\left[1 - \Phi\left(a_t^*(B) + \xi_t \Delta + \frac{1}{2}\right)\right] a_t^*(B), \left[2 - \Phi\left(a_t^*(N) + \xi_t \Delta\right) - \Phi\left(a_t^*(N) + \xi_t \Delta + 1\right)\right] a_t^*(N)\right\}$$

Furthermore, we show that when the influencer has higher reputation of being H type, both his horizontal content design x_t and the amount of advertisements a_t increases monotonically, as summarized in the Lemma below.

Lemma 2. *There exists a threshold $\bar{\theta}$, such that when $\theta_t < \bar{\theta}$, the influencer chooses a niche content design ($x_t = 0$); when $\theta_t \ge \bar{\theta}$, the influencer chooses a broad content design ($x_t = \frac{1}{2}$). The amount of* *advertisements a_t in equilibrium increases with θ_t monotonically.*

5.3 Accumulation of Past Subscribers: $\gamma < 1$

Now we turn to a more realistic situation that the depreciation rate $\gamma < 1$. This reflects that some of the past audiences ($1-\gamma$ fraction) accumulates to future periods. First, I show that an influencer prefers to design his contents such that they always (weakly) favor one type of audiences. In Lemma 3, this result allows us to therefore restrict the choice of x_t to $[0, \frac{1}{2}]$.

Lemma 3. *For any history of content design, $H_t = \{(x_0, a_0), (x_1, a_1), \dots, (x_{t-1}, a_{t-1})\}$, denote $\Pi_s(H_t)$ as the influencer's profit at period $s \leq t-1$ under such history. Consider another content design history $\tilde{H}_t = \{(\tilde{x}_0, a_0), (\tilde{x}_1, a_1), \dots, (\tilde{x}_{t-1}, a_{t-1})\}$, such that $\tilde{x}_s = x_s$ if $x_s \leq \frac{1}{2}$ and $\tilde{x}_s = 1 - x_s$ if $x_s > \frac{1}{2}$, $\forall s \leq t-1$. Then $\Pi_s(\tilde{H}_t) \geq \Pi_s(H_t), \forall s \leq t-1$.*

The intuition is that with the presence of past subscribers, the content design is history dependent. If an influencer accumulates more type A audiences in the past, a content design that favors type B does more harm to the type A past subscribers than the harm to type B subscribers with a reversed favoring type A content. Therefore the influencer is always weakly better off to consistently prefer one type of audiences when the content design is not the most broad. For tractability, from this section on I assume that $\phi(\cdot)$ is not too skewed. More specifically, I make the following assumption:

Assumption 1. $(1+\gamma)\phi(\mu - \frac{1}{2}) \geq \phi(\mu)$

In the analysis of the model, this assumption serves as a sufficient condition, such that no matter how drastic the reputation increment is, the influencer will not abandon any of his past audience on the equilibrium path.[19] We first show that when θ_0 is low enough, the influencer starts from a niche design. Within the niche design periods the amount of advertisements strictly increases. Denote $\underline{u}_t^j$ as the lower bound of type $j \in \{A, B\}$ new subscribers' match value. We also show that $\underline{u}_t^j$ is weakly decreasing for both types during the niche periods.

[19] I provide a numerical simulation where Assumption 1 is relaxed in Appendix C, and find that the optimal content design strategy is robust.

Lemma 4. *Compared to the $\gamma = 1$ situation, when $\gamma < 1$, the influencer stays continuously producing the most niche content for longer periods of time. While the influencer provides niche content, the amount of advertisements monotonically increases, and the lower bound of subscribers' match value to the influencer monotonically decreases.*

When $\gamma < 1$, at every period the demand that the influencer faces has two components. The first component is the old subscribers that are carried out from the past periods. The distribution of each period's subscribers is a truncated log-concave distribution for both types. The second component is the new audiences that is log-concave distributed and of mass 1 for each type. Note that at $t = 0$, the problem for any γ is the same. So if under the $\gamma = 1$ situation, the influencer starts from a niche content, for any $\gamma < 1$, the influencer also designs a niche content. At $t = 1$, the past subscribers are type A leaned because they are truncated by a niche content in $t = 0$ that favors A audiences. The fact that the majority of past subscribers favor topic A provides excessive incentive for the influencer to choose a niche content that favors A. Therefore, as long as the $\gamma = 1$ influencer chooses niche content, the $\gamma < 1$ influencer also chooses niche content, since he has the aligned incentives to provide niche content for both the old subscribers and the new potential audiences.

Next, we analyze the situation when the influencer accumulates high enough reputation, and (possibly) the content design moves broad. I show that when there exists old subscribers ($\gamma < 1$), the bang-bang content design strategy in Lemma 1 may not hold: an influencer may design a content that is neither the most niche ($x_t = 0$) nor the most broad ($x_t = \frac{1}{2}$). However, the direction of content design changing remains the same as in Lemma 1. That is, the content design monotonically becomes broad and advertisements monotonically grows when θ_t increases over time.

Proposition 1. *The horizontal design for contents, x_t, increases monotonically over time and has an upper bounded of $\frac{1}{2}$. That is, the horizontal content design monotonically moves from niche to broad. The amount of advertisement, a_t, also increases monotonically with t.*

Proposition 1 fits our observation from the data in Section 4.1 and 4.2 that topic dispersion and the extent of advertising grow over time. The intuition for the Proposition 1 is as follows.

An influencer faces two different groups of demand: one is the newly arriving audiences that are symmetric between both types, and the other is his old subscribers, where type A subscribers are more than type B. When an influencer accumulates high enough reputation to be H type and only faces the new audiences, he should design the most broad content. However, such a drastic change would harm his old type A subscribers and make them stop consuming his content. Therefore, the influencer should trade-off between monetizing from newly arrival audiences or from old subscribers when designing x_t. This drives the influencer to design a gradually broadening content such that the old type A subscribers will not be discarded while still pleasing the type B new audiences. Over time type A old subscribers depreciate faster than type B, and are replaced by the new audiences who are more balanced between the two types. This further spurs the influencer to design a more broad content. Two different forces drive an influencer to provide more advertisements. First, when an influencer accumulates higher reputation, his audiences have higher expectation for his contents and therefore can bear more advertising posts. Second, an influencer gains more subscribers over time, and this increases the benefit from monetizing the existing subscribers, rather than reducing advertisements to attract more new audiences.

6 Further Implications and Empirical Evidence

The theoretic framework makes predictions on the equilibrium path that fit with the stylized facts that are established in Section 4, namely, the topic diversity and the extent of advertising grow over time. In this section, I derive three further implications based on the theoretic framework, and provide further empirical evidence from the data. First, I examine how influencers react to a reputation shock, where they are allowed to label an "original tag" to their blog posts. Second, I examine how an exogenous shock on audiences' preference changes the content design of influencers at different stages of their life cycle. Finally, I study how the distribution of ratings evolve over a new influencer's life cycle.

6.1 Exogenous Reputation Shock: Original Tag

Two key driving forces, the growth in reputation and the accumulation of subscribers, lead to the dynamics of content design in the baseline model. We cannot directly observe reputation of influencers, since in the model reputation co-moves with time: the longer an influencer survives in the market, the higher reputation he gains as those influencers that have proven to be incompetent have already left the market. However, other unobserved factors may also change over time. For example, influencers gradually learn how to write better articles, or gradually establish better relationship with advertising agencies. These may also lead to more diverse topics or more advertising posts. This brings a challenge to the identification of our mechanism: can we find an exogenous shock to the reputation of influencers to cross-validate the model?

I use the ”original tag” function of Wechat Official Account platform as a reputation shock. First launched in Jan. 22, 2015, Wechat allows some official accounts to tag their blog posts as ”original” to distinguish original articles from articles of other sources (for example, forwarded articles from other media or official accounts). This function is not by default accessible to all the influencers. Instead, Wechat Official Account platform sends invitations to official accounts that they believe to be of ”high quality”, and the standard for sending invitations is not transparent.[20] Therefore, the timing of the arrival of this invitation can be seen as a relatively exogenous shock that is unexpected to the influencer. It is also a pure reputation shock: an article with original tag signals that the quality of that influencer is recognized by Wechat official, which boosts his reputation among readers. But the arrival of original tag invitation does not change other unobserved factors, like the influencer's ability to write posts, or his relationship with advertising agencies, etc.

I first present some preliminary evidence that the ”original tag” increases topic diversity and advertising in blog posts by the event study method, which is commonly used in literature to identify ”badge effects” of reputation on online platforms (for example, Hui et al., 2016, Cheng et al., 2020) In a nutshell, I compare short periods of time (6 months or 3 months) for each in-

[20]Wechat provides several broad standards. E.g., less violation records, subscriber base, the quality of contents, etc. Online discussion shows that many practitioners are still confused and question the standards that Wechat provides. For example, see https://www.douban.com/group/topic/88550314/

fluencer before and after the arrival of the original tag invitation, which is identified as the first time I observe a blog post with original tag for each influencer. The econometric specification is as follows:

$$y_{it} = \alpha_i + \alpha_k + \beta_1 PostOriginal_t + \beta_2 Time_t + \beta_3 InitialDate_i + \beta_4 OriginalTime_i + \epsilon_{it}$$

for influencer i, time t and main topic k. y_{it} is either $ln(Dispersion_{it})$ (as defined in Section 4.1) or $ln(AdvShare_{it})$ (as defined in Section 4.2). α_i is the influencer fixed effect and α_j is the fixed effect for the major topic k of all the publications of influencer i at time t, $\mathcal{A}_{it}$. Time $t = 0, 1, 2...11$ is the relative time (in months) to the original tag shock for each influencer. $t = 0$ is 6 months prior to the original tag shock and $t = 11$ is 6 months post the shock. $PostOriginal_t$ is a dummy variable that equals to 1 if $t >= 6$. $InitialDate_i$ is the time of influencer i's first publication and $OriginalTime_i$ is the time that I first observe blog post with original tag for influencer i.

Table V presents the main result. Column (1) - (4) is the effect of an influencer getting original shock on the topics dispersion in the contents. Column (5) - (6) is the impact on the share of advertising posts. For column (1) and (2), I use the average Euclidean distance for all the publications of influencer i, time t as a measure of dispersion. For column (3) and (4), I use the standard deviation of publications in the topics space instead as the dispersion measure. In column (1), (3) and (5), I uses 6 months data for both the pre-trend and post-trend. And in column (2), (4) and (6), the pre and post trends are 3 months. The economic magnitude of the reputation shock brought by the original tag is sizeable. Getting the original tag increases 3.6-7.1% of topics dispersion in contents, which equivalents to about 5.2-10.1 months of the increment in topics dispersion. It also increases 43.2-53.7% advertising posts, which equivalents to 7.8-11.3 months of the increasing in advertising posts.

The underlying assumption for the event study method is that the arrival of original shock is purely random, which is potentially problematic in our context. The original tag function is awarded to official accounts that are seen as high quality by Wechat official. Therefore it is possible that Wechat is more likely to select accounts whose reputation is going to rise shortly and we wrongly attribute such a rise to the impact of original tags. For influencer i, let the original tag happens at time $\hat{t}_i$. The sample average treatment effect of the treated (SATT) of the original tag should be $\lambda = \mathbb{E}(y^1_{i,\hat{t}_i+s} - y^0_{i,\hat{t}_i+s})$. The superscript $OT_{it} = \{0, 1\}$ denotes whether the account receives original tag function, and y represents either the topic diversity or advertising share. The fundamental problem is $y^0_{i,\hat{t}+s}$ is not observable after an account i acquires the original tag function, and we do not have a fully identical influencer account i' that has the same history as influencer i, but does not acquire original tag function at time $\hat{t}_i$.

The idea to solve the problem is to identify close matches among all the potential candidates who have not received the original tag shock, and then we could apply a difference-in-difference (DID) approach to identify the treatment effect of original tag shock. This is a popular approach in the literature of reputation badges to alleviate endogeneity concerns (for example, Elfenbein et al., 2012, Elfenbein et al., 2015, Cheng et al., 2020, Tripathi and Kyriakou, 2020, etc.). I use a nonparametric matching approach, "coarsened exact matching" (Iacus et al. 2012, Blackwell et al. 2009), which has been shown to be better in balancing pre-treatment covariates and more efficient compared to propensity score matching (King and Nielsen, 2019).[21]

The selection of controls proceed as follows. First, I select all the observations of influencer j at age t_j, such that within one year around t_j, $[t_j - 6, t_j + 5]$, the influencer j does not receive the original tag invitation, and j is older than 3 months ($t_j \geq 3$). The collection of (j, t_j) that satisfies those conditions are potential controls, and all the $(i, \hat{t}_i)$ are in the treatment group. Second, I choose a relatively small set of covariates, which we would like to match the pretrends. I choose the average clicks, the average likes, and the standard deviation of content topics. Third, I create a large number of strata which covers the entire support of the joint distribution of those covariates. I use the default strata segregation method in (Blackwell et al., 2009). Next, each observation in the treatment group or the potential control is allocated to a unique stratum, and I drop all the strata that does not have at least one observation from potential control group and one from the treatment group. Last, for each stratum, I use all the observation from the potential control group as the matched control for treatments in this stratum, and I put equal weights to each control. Last, for each (j, t_j) that in the control group, I use one year data around t_j, $(j, [t_j - 6, t_j + 5])$ for a

[21] As a robustness check, we provide an estimation result using propensity score matching method in Appendix C, and the estimation result does not qualitatively change.

standard panel data DID approach.

Figure V and VI shows the time trend of the average content dispersion (measured by average Euclidean distance of $\mathcal{A}_{it}$) and share of advertising posts, among the treatment group and the CEM matched control group. Despite close pre-trend between the treatment and control groups, the treatment group spikes rapidly in both content dispersion and advertisement shares, relative to the control group. Table VI shows the statistical comparison between the unmatched potential controls and the CEM matched control group. It can be seen that not only the difference between the means is smaller after the matching (reflected both in the mean and the t-statistics), but the distributions of the treatment and control are closer to each other under the CEM matching. The DID econometric specification is as follows:

$$Y_{it} = \alpha_i + \alpha_k + \beta_1 Post_{it} + \beta_2 Treat_i + \beta_3 Post_{it} \times Treat_i + \beta_4 Age_{it} + \beta_5 InitialDate_i + \epsilon_{it}$$

for influencer i, time t and main topic k. y_{it} is either the topic diversity, $ln(Dispersion_{it})$ or portion of advertising posts, $ln(AdvShare_{it})$. $Post_{it}$ is a dummy variable that equals to 1, if influencer i is in the treatment group and the time is after the first period he publishes an original tag content $\hat{t}_i$, or if (j, t_j) is in the control group and $t \geq t_j$. $Treat_i$ is a dummy variable that equals to 1 for treatment group observations. Age_{it} is the age (in month) of influencer i at time t, and we control for the influencer i fixed effect by α_i and main topic k fixed effect for publications in $\mathcal{A}_{it}$.

Table X shows the result. Same to the event study table, columns (1) and (2) uses average Euclidean distance as the measure of topic diversity, while (3) and (4) uses the standard deviation. Columns (1), (3) and (5) use 6 months of pre-trend and post-trend, and (2), (4) and (6) use 3 months of data. I find that compared to the control group, the original tag shock increases 5.2% to 6.5% of content dispersion, which equivalents to the increment of topic dispersion of 7.5 - 9.3 months. On the share of advertisements, receiving the original tag function increases the amount of advertisement by 65% to 66%, which equivalents to about 14 months of the increment in advertising posts.

It is worth discussing some possible caveats of this empirical application. First, the impact of this reputation shock may be underestimated on both content diversity and advertising, as official accounts can only give original tags to blog posts they write on their own. When an influencer gets the original tag function, he has greater marginal benefit to write original articles to show to his audiences that he got recognized by Wechat Official Account platform, which are likely to be around the core topic that the influencer is most knowledgeable in. This would reduce the relative amount of repost articles from other fields (which contribute to the topics diversity) and sponsored advertisements.

Another possible source of bias is the anticipation effect of the treatment. Although the exact standard to be invited is not transparent, it is still plausible that influencers understand that when they become more established (more subscribers, higher average clicks and likes, etc.) they will be more likely to receive the original tag invitation and their reputation is likely to be boosted in the near future. Therefore they may strategically adjust their contents *before* receiving the shock if the influencers are forward looking. Such bias would underestimate the impact on content diversity, and overestimate the impact on advertising share. This is because if an influencer expects a sudden increment in reputation in the near future, he is going to target audiences with more diverse preferences. And it is beneficial to adjust the content design to the broader side prior to the shock so that the base of subscribers on different topics become more balanced distributed. On the other hand, if the influencer expects an increase in reputation in the near future, the marginal benefit of having more subscribers is greater since he expects to post more advertisements after he receives the original tag invitation. Therefore the influencer has stronger incentive to reduce the advertising level prior to the shock if he anticipates the shock.

6.2 Exogenous Preference Shock: the Arrest of Meng Wanzhou

Audiences typically have persistent preference on contents. But sometimes their preference may be disrupted due to important new events that affect the whole society. For example, Covid-19 has widespread impact to everyone since it emerged. Should an influencer, especially a young one allocate some routine contents into it to attract audiences? And how to trade off between the new event and the original content routine? Additionally, an unexpected event brings an exogenous shock to the preferences of the audiences, which provides a cross-validation to the theoretical model. In this section, I provide an extension to the baseline model that incorporates the new event shock, and deliver empirical evidence from an important shock in Chinese society during my data period of time: the arrest of Meng Wanzhou.

Assume that an unexpected new event emerges at time t and disappears the next period. In all other periods except t, the game is the same as the baseline model. At time t, an influencer could allocate a share of his content, $s_t^N \in [0,1]$, to the new event and keep the rest $1 - s_t^N$ fraction of the content to his routine contents. I assume that all types of consumers have some interests to the new event for its wide spread impact. That is, a fraction s_t^N of content on the new event gives both types of audiences a utility $u_N(s_t^N)$, and $u_N(\cdot)$ is assumed to be increasing, strictly concave and twice-differentiable over $[0,1]$. For the rest $1 - s_t^N$ part of routine contents, the influencer still designs the horizontal position x_t, and the extent of advertising a_t. But audiences' payoffs from the routine content are depreciated proportionally to the fraction of remaining routine content. To sum up, the payoff that audiences gain from consuming the content at time t is:

$$W_{it} = \begin{cases} (1 - s_t^N)(u_i - x_t) + u_N(s_t^N) - \mathbb{1}\{\text{Bad}\}\Delta - a_t, & \text{if type A} \\ (1 - s_t^N)(u_i - (1 - x_t)) + u_N(s_t^N) - \mathbb{1}\{\text{Bad}\}\Delta - a_t, & \text{if type B} \end{cases}$$

The influencer chooses optimal content design (x_t^*, a_t^*, s_t^{N*}) at time t, given the accumulation of past subscribers of both types audiences, denoted by $F_t^A(u_i)$ and $F_t^B(u_i)$. We have the following proposition:

Proposition 2. *For an influencer, the optimal allocation to the new event in his contents increases over time. That is,* $\forall t' < t$, $s_{t'}^{N*} < s_t^{N*}$.

The intuition behind Proposition 2 is such that the influencer trades off between the importance of the new event and its match to his audiences. For a young influencer, the majority of his past subscribers favors one topic (w.l.o.g type A), and his routine content design is also specialized in topic A. Although the new event might be important to both types audiences, he does not gain much from its benefit to type B audiences, as no much type B audiences has subscribed to him, and he does not plan to attract type B audiences as well. Therefore, the relative gain from allocating content to the new event is low for young influencers. For an old influencer, his past subscribers are more balanced distributed between two types, and his routine horizontal design is also a broad design. Discussion on the new event is therefore more aligned with his horizontal content design incentive compared to young ones. We have the following prediction for an empirical test:

Prediction 1. *When a new general interest event arises in the society, old influencers allocate more content to this event compared to young influencers.*

To empirically examine different reactions of influencers at different stage of life cycle, I take advantage of the arrest of Meng Wanzhou to test this prediction. On Dec 1, 2018, the CFO of Huawei, Meng Wanzhou, get arrested in the Vancouver International airport at the request of the United States.[22] There are several salient advantages for using Meng Wanzhou detention case to test this prediction. First, this case is an unexpected shock with a clear starting time. No influencer could have anticipated this event. Second, due to the dominant position of Huawei in the high-tech industry of China, this event had deep and wide impact to all of Chinese society and has received tremendous attention from the general public. Numerous followed up articles in many fields relate Meng Wanzhou to their fields (for example, fashion, real-estate, relationship, education, etc.). Third, the arrest of Meng Wanzhou is close to the end time of our data. This allows us to use the data from most influencers in the dataset to carry out the empirical test. Fourth, compared to events with multiple keywords and cannot be easily categorized (like the China-US trade war), there is clear way to label Meng Wanzhou related articles - her name. Because prior to this shock, there is almost no article contains Meng's name (see Figure VII). And lastly, unlike some other big events like the beginning stage of Covid-19, there is no sign that Meng Wanzhou related articles were subject to censorships from Wechat. This reduces the concern that old influencers may know censorship standards better, or Wechat was being more lenient with old influencers.

To identify Meng Wanzhou related articles, I search for key word ”孟晚舟” (Meng Wanzhou) in all articles. Any article that contains Meng's name is labelled as Meng Wanzhou related. I apply a diff-in-diff approach to estimate the different reactions to the Meng Wanzhou detention between

[22]See https://en.wikipedia.org/wiki/Arrest_of_Meng_Wanzhou for a detailed description of the event.

the old and new influencers. More specifically, I estimate the following regression:

$$MengShare_{it} = \alpha + \beta_1 New_i + \beta_2 Post_t + \beta_3 New_i \times Post_t + \epsilon_{it}$$

where i is influencer and t is days. $MengShare_{it}$ is the share of Meng Wanzhou related article in influencer i day t posts. New_i is a dummy variable that equals to 1 for all influencers that are born within 1 year of Meng Wanzhou Detention (after Dec. 1, 2018), and before the starting time of pre-trend (Nov. 1, 2019). $Post_t$ is a dummy variable that equals to 1 if the time is post Meng Wanzhou detention. I assume that the pre-trend and post-trend are both 30 days.

Figure VII shows the time trend of the average Meng Wanzhou related articles fraction among the old and the new influencers. It shows that the old official accounts allocate more content to Meng Wanzhou event. Table VIII shows the DID regression result. The column (1) is the main result. To interpret the economic magnitude, the old influencers publish about $\frac{0.0123}{0.0123-0.003} - 1 =$ 32.3% more Meng Wanzhou related articles than the young influencers in fraction of their publication. I provide several robustness checks. Topic 22 contains high weight keyword "Huawei" (see Table II). In column (2), I rule out all the accounts that the topic mode of whose publications is topic 22. In column (3) I drop influencers that publish most in topic 5 as well, since it is the topic that is economics/industry/market focused. In column (4), I use both "Meng Wanzhou" and "Ren Zhengfei" (Meng's father and the CEO of Huawei) as the keywords to label Meng Wanzhou arrest related articles. In all the variations, we observe a significant negative estimation of the coefficient of $Post_i \times New_t$ as predicted by the theoretical model, and the economic magnitudes are similar: the old accounts publish 33.6% to 38.2% more Meng Wanzhou related articles in fraction than the young ones.

6.3 Stochastic Quality, Spreading and Rating

Most social platforms allow users to express their feelings towards a content through ratings. For example, "thumb up/thumb down" on YouTube, "like" on Facebook and Wechat Officail Account, or "heart" on Instagram. In this section, I provide an extension to the baseline model that aims to understand the pattern of dynamics in the ratings of the contents. A gap between my baseline model and reality is that in my baseline model, the quality and the number of views of an influencer's contents are deterministic and evolves smoothly over time. However, in reality, we observe variation in both the ratings and the number of views for contents an influencer produces within a short period of time.[23] To capture this, I consider a scenario where there exists stochastic shocks to the quality of contents on top of the content design $\pi_t = (x_t, a_t)$. The readers of a post induce word of mouth to spread the content to outsiders based on the quality of the content, and all audiences rate the content by reporting the subjective experienced utility.

I assume that there exists a quality shock, ϵ_t that is common to all the audiences with c.d.f $F(\epsilon)$. ϵ has support over $[\underline{\epsilon}, \bar{\epsilon}]$ and w.l.o.g $\mathbb{E}(\epsilon) = 0$. The realization of ϵ_t is independent of content design π_t, and the quality shock realization is only observed by the audiences after they consume it. Instead of an exogenous group of new audiences that get access to the influencer's content at each period, we now assume that the existing subscribers share content through their social media and spread it to their friends based on the realization of the quality shock. A higher quality shock (ϵ) content has higher chance of being shared by any subscriber and subsequent readers. To capture this, we assume that the average number of outside readers that each subscriber reaches through the word of mouth network is $G(\epsilon)$, which is increasing in ϵ. Compared to the subscriber, the tastes of her connected outsiders are closer to those of the general public. More specifically, assume that $\alpha \in [0, 1)$ fraction of her connected outsiders have the same taste as the subscriber (same type and same idiosyncratic match value), and $1 - \alpha$ fraction are randomly drawn from the general public. Intuitively, $\alpha = 0$ is the extreme case that all agents connect to random outsiders in the network (e.g., an anonymous forum), and a greater α means higher similarity among the linked audiences in the network.

Upon reading the content, each audience rate the content by reporting their subjectively experienced utility.[24] Denote by R_{it} the audience i's rating at time t under content design (x_t, a_t) and

[23] For example, the average coefficient variance of clicks in the Wechat Official Accounts data (which I use later for empirical tests) is 0.802, and the average coefficient variance of likes is 1.41 at the influencer-month level.

[24] This assumption is aligned with literature like Bondi (2019) and Brandes et al. (2019).

the quality shock ϵ_t, we have:

$$R_{it} = \begin{cases} u_i - x_t - \mathbb{1}\{\text{Bad}\}\Delta - a_t + \epsilon_t, & \text{if type A} \\ u_i - (1 - x_t) - \mathbb{1}\{\text{Bad}\}\Delta - a_t + \epsilon_t, & \text{if type B} \end{cases} \tag{4}$$

Let $F_t^A(u_i)$ and $F_t^B(u_i)$ be the cumulative distribution functions of the existing subscribers that are of type A and B respectively. Under the content design (x_t, a_t), the lower bounds of each type subscribers that would consume the content are $\underline{u}_t^A = a_t + \xi_t\Delta_t + x_t$ and $\underline{u}_t^B = a_t + \xi_t\Delta_t + 1 - x_t$ respectively. We denote $Q_t = \mu - a_t - \frac{1}{2} + \epsilon_t$ as the *unbiased quality* of the content at time t, which is the average rating if the content is rated purely by the general public. Denote $\bar{R}^j(F_t^j) \int_{\underline{u}_t^A}^{\infty}(u_i - a_t - x_t + \epsilon_t)dF_t^A(u_i), j = A, B$ be the average rating of type $j = A, B$ subscribers. Then the mean of rating across all readers, $\mathbb{E}(R_{it})$, is denoted by:

$$\mathbb{E}(R_{it}) = \frac{(1 + \alpha G(\epsilon_t))[\bar{R}^A(F_t^A) + \bar{R}^B(F_t^B)] + (1 - \alpha)\left[\int_{\underline{u}_t^A}^{\infty} dF_t^A(u_i) + \int_{\underline{u}_t^B}^{\infty} dF_t^B(u_i)\right] G(\epsilon_t)Q_t}{\left[\int_{\underline{u}_t^A}^{\infty} dF_t^A(u_i) + \int_{\underline{u}_t^B}^{\infty} dF_t^B(u_i)\right](1 + G(\epsilon_t))} \tag{5}$$

which is the weighted average between the rating of existing subscribers and the outsiders. First, note that given the content design π_t, a higher quality shock ϵ_t leads to a higher *unbiased quality* Q_t, and more outside readers as $G(\epsilon)$ is increasing in ϵ. Second, $\mathbb{E}(R_{it}) > Q_t$. This bias comes from the selection effect, such that the past subscribers of the influencer have higher average match value with the influencer than the general public does. Lastly, denote by $\Delta_t(\epsilon_t | x_t, a_t, F_t^A, F_t^B, \alpha) = \mathbb{E}(R_{it}) - Q_t$ the *selection bias*. Δ_t is decreasing in ϵ_t, and if $G(\epsilon)|_{\epsilon\to\infty} \to \infty$ and $\alpha = 0$, $\Delta_t|_{\epsilon_t\to\infty} \to 0$. Intuitively, when a content has a very high quality shock such that most of its readers are from the general public, the selection bias from subscribers diminishes, and the overall rating converges to the *unbiased quality* Q_t. Interestingly, the fact that worse contents are more over-rated through the stronger selection bias also implies that a better quality content may have even lower rating than a worse content.

Next, we examine the dynamics of ratings for influencers. Here we take the content design dynamics in the Proposition 1 as given, such that the amount of advertisements grows over an influencer's life cycle, and the horizontal design x_t monotonically increases. The following proposition shows that both the true quality and the selection bias monotonically decreases over time, which results a monotonically decreasing trend on ratings in expectation.

Proposition 3. *For any $t < t'$ in an influencer's life cycle, for any fixed quality shock realization ϵ, and for any $\alpha \in [0, 1)$, we have:*

1. *The unbiased quality Q_t monotonically decreases: $Q_t(\epsilon) \geq Q_{t'}(\epsilon)$.*
2. *The selection bias monotonically decreases: $\Delta_t(\epsilon) \geq \Delta_{t'}(\epsilon)$.*
3. *The difference in selection biases for better versus worse contents decrease over time: $\Delta_t(\epsilon') - \Delta_t(\epsilon) \geq \Delta_{t'}(\epsilon') - \Delta_{t'}(\epsilon), \forall \epsilon < \epsilon'$.*

To see the intuition, note that the downgrading of *unbiased quality* Q_t is directly driven by the increasing amount of advertisements a_t over time as in Proposition 1. In my model, because young influencers do not have reputation for producing good contents, only those audiences who have high match value with him and those who likes the topic he specializes in would subscribe to him. Both the initial specialization of contents and the high selection in match value increases the selection bias for young influencers. Over time, the influencer gains reputation for producing good contents, and his contents become broader. This drives lower match valued audiences on both types to subscribe to the influencer, which reduces the selection bias among his subscribers.

The last point in Proposition 3 shows an interesting asymmetry between the low quality and high quality articles. For younger influencers, the rating is more biased in favor of low quality articles than the high quality ones. The difference in the biases can be large enough such that a higher rated content actually have a lower quality. Over time, the difference in the biases, although never fully erased, is attenuated. This implies that both the audiences and the platform should be especially careful when trying to refer quality from ratings for the *young* influencers.

To test the Proposition 3, note that ϵ_t is identically distributed across periods, and a higher ϵ_t corresponds to a higher views of the content at time t conditional on content design π_t and reputation θ_t. Let the collection of all the contents produced around time t is $\mathcal{C}_t$, and let the content that is at quantile q in views of $\mathcal{C}_t$ be $c_t(q)$. For any quantile q, the quality shock ϵ_t that corresponds to $c_t(q)$ are the same across all periods of time. I take advantage of that to derive the following empirical prediction:

Prediction 2. *For an influencer i, let the average rating for q-quantile contents at time t be* $\mathcal{R}_{it}(q)$*. For any* $q \in [0,1]$*, the average rating for q-quantile contents decreases over time, that is,* $\frac{\partial \mathcal{R}_{it}(q)}{\partial t} < 0$*. And the rating of lower quantile contents decreases faster:* $\forall q < q'$, $\frac{\partial \mathcal{R}_{it}(q)}{\partial t} < \frac{\partial \mathcal{R}_{it}(q')}{\partial t}$.

Empirically, my data contains the counts of likes and clicks in about 3 days after the article got published, and I use the likes-clicks ratio as a proxy of rating. Before the empirical analysis, it is worth discussing several potential problems with online rating under the Wechat data context. The first problem is fraud rating (Aral, 2014, Luca and Zervas, 2016). This problem is relatively minor on Wechat compared to other platforms, because Wechat requires a valid cellphone number to register a normal user account. Besides, a normal user can contribute at most 5 clicks counts per day for each article, which reduces fraud clicks. The second is rating inflation, such that the distribution of online ratings gradually move left skewed, and form the so-called "J-shape" in the literature (Filippas et al., 2019, Hu et al., 2009). My theoretical model predicts that over time average rating *decreases*. So if rating inflation does exist on Wechat, such bias would only underestimate the significance of the estimation. Lastly, the observation in my data for both clicks counts and likes counts are truncated at 100,000. This is because Wechat only displays "100,000+" when the clicks or likes are above 100,000. About 15% of the total articles in my data have 100,000+ clicks. To solve this problem, I select the official accounts (563/1002) that have relatively few articles (less than 5%) with 100,000+ clicks, which are affected less by the data truncation problem. Among them, I calculate the average likes-clicks ratio of articles that have 90,000 - 100,000 clicks, at influencer i age t (in month) level. And I use it as the proxy of the likes-clicks ratio of 100,000+ clicks articles for influencer i at time t.

Figure VIII shows the time trend of average likes-clicks ratio for articles at different quantiles, where the x-axis represents the age of influencer (in months). Note that the main driving force of the dynamic in rating comes from two sources: the revealing of true quality of the influencer and the broadening of the content topics. Both happens in the relatively early stage (the growing stage) of an influencer's life cycle. The econometric specification is as follows:

$$ln(LCRatio_{it}) = \alpha_i + \alpha_k + \beta_1 Age_{it} + \beta_2 InitialDate_i + \epsilon_{it}$$

where α_i is the influencer i's fixed effect, and α_k is the fixed effect of the major topic in $\mathcal{A}_{it}$. Table IX shows the result. Column (1) is the regression result that using the total sample; column (2) uses the 0 - 25% quantile articles in clicks (low clicks); column (3) is the result that uses 25 - 75% quantile articles (middle clicks); and column (4) is the result of 75 - 100% quantile articles (high clicks). As predicted by the model, all the estimations of β_1 are significantly negative, and we have $\beta_1^{low} < \beta_1^{middle} < \beta_1^{high}$. The differences between β_1^{high} and β_1^{middle} or β_1^{low} are also statistically significant (p=0.001). Column (5) - (8) uses the data in the beginning 2 years of each influencer, as influencers grow most in this period and should have greater change in the distribution of ratings. The results echo with this intuition: the average decreasing rates doubled in the beginning two years compared to the whole range of data. To interpret the economic magnitude, in the beginning 2 years, the overall average rating drops about 7.2% per year. The rating decreasing rate for the top quarter of articles is 1.6% per year, while the decreasing rate for the bottom quarter articles is 13%.

Last, it is worth noting that a big literature in economics highlights the importance of learning by doing (for example, Arrow, 1971, Freeman and Soete, 1997). The key insight is that over time, bloggers may learn how to produce better contents and they can allocate more contents to other topics and advertisements. However, the downward trend of rating suggests that the learning by doing mechanism may not be the first order driving force in content dynamics. Under the learning by doing mechanism, the unbiased quality of contents should be increasing over time.[25] The two forces in the rating dynamics make the opposite predictions: learning by doing leads to an upward trend in rating, and selection bias leads to a downtrend in rating. Figure VIII and Table IX suggests that even among the top quarter quantile of articles in clicks (where the selection bias is minimal), the overall rating still trends down. This implies that even if the learning by doing effect does exist, it is relatively small compared to changes in the selection bias.

[25] This is a classical result under mild regulatory conditions, log-concave distribution of consumers' match value.

7 Conclusion

This paper studies the dynamics of content design both theoretically and empirically with influencer data of Wechat. There are two sides of the content design problem. On the horizontal side, a content can be designed to be niche and attract audiences with specific horizontal taste, or a can be designed broad such that audiences with wider ranged tastes would be attracted. On the vertical side, an influencer has to trade-off between monetizing the current audiences (including new audiences and the past subscribers) with more advertisements and lower quality, or gaining more new audiences through a higher quality content.

I provide a theoretic model to capture the changes in the incentives on these two sides over an influencer's life cycle. I argue that two forces drive the changes in the content design: reputation of providing good contents, and the accumulation of the old subscribers. In the early stage of an influencer career, he has not yet established reputation for providing good contents. Audiences bear higher risk to consume his content and subscribe him, and specialization in a niche topic at least gains him some audiences on that topic. But for established influencers, such incentive is reversed, as audiences are more certain that he provides good contents, and the audience broaden the content to reduce the down-side risk. This niche to broad mechanism can be attenuated by the accumulation of old subscribers. When it takes too long for an influencers to establish reputation, he might have already gained a large population of subscribers who are interested in niche topics. This gives him excess incentive to seed high amount of advertisements in his contents and stay niche, just keeping the current size of subscribers. A higher reputation gives more room for the influencer to seed advertisements, and the growing size of subscribers also provides the influencer a higher incentive to monetize from the existing subscribers.

Rating is crucial in evaluating the quality of contents on most influencer platforms. This research shows that solely rating is not sufficient and is biased in evaluating quality, as young influencers and old influencers are structurally different from each other. A young influencer has niche design in content and his subscribers are concentrated on that niche topic. Therefore young influencers selects more subscribers who love him, which causes upside bias in rating compared to the average view of total population. Same bias happens as well on low viewed contents, as those contents are more viewed by self-selected subscribers rather than the outsiders. This raises alert that even if two contents have similar ratings, the true qualities of the two contents can be very different. The age of the influencer and the number of views are also critical in evaluating a content.

On the empirical side, I use data from Wechat Official Account platform and study the historical publications of top influencers. I show that data confirms my theoretical predictions quite well: topic covering and advertisements grows over time; a positive reputation shock increases both content diversity and advertising posts; average rating of contents decreases and the ratings of high and low clicked contents diverge; and old influencers react to an event that changes the preference of audiences, the arrest of Meng Wanzhou, more than the young influencers.

This paper may lead to several potential future researches, both theoretically and empirically. On the theoretical side, this research has mainly focused on the early stage of a new influencer's life cycle. Data suggests that after reaching the maturity stage, some influencers gradually lose attentions from subscribers and decline. It would be helpful for us to understand the mechanism behind such a declining stage, so that we can have a full picture of the whole life cycle of influencers. Empirically, this data may help us to understand other features in the influencer market, for example, click baits and fake news. Recent theoretical literature (like Deb et al., 2020) studies the incentive of media providing fake news, and has made predictions on the pattern of fake news. It would be interesting to test these theoretical frameworks with this new dataset.

References

Philippe Aghion, Reinhilde Veugelers, and Clément Serre. Cold start for the green innovation machine. Technical report, Bruegel policy contribution, 2009.

Maaz Amjad, I Voronkov, A Saenko, and A Gelbukh. Comparison of text classification methods using deep learning neural networks. In *Proceedings of the 20th International Conference on Computational Linguistics and Intelligent Text Processing (CICLing)*, 2019.

Simon P Anderson and Régis Renault. Comparative advertising: disclosing horizontal match information. *The RAND Journal of Economics*, 40(3):558–581, 2009.

Sinan Aral. The problem with online ratings. *MIT Sloan Management Review*, 55(2):47, 2014.

Kenneth Joseph Arrow. The economic implications of learning by doing. In *Readings in the Theory of Growth*, pages 131–149. Springer, 1971.

Jie Bai, Maggie Chen, Jin Liu, and Daniel Yi Xu. Search and information frictions on global e-commerce platforms: Evidence from aliexpress. Technical report, National Bureau of Economic Research, 2020.

Joe S Bain. Relation of profit rate to industry concentration: American manufacturing, 1936–1940. *The Quarterly Journal of Economics*, 65(3):293–324, 1951.

Heski Bar-Isaac, Guillermo Caruana, and Vicente Cuñat. Search, design, and market structure. *American Economic Review*, 102(2):1140–60, 2012.

Heski Bar-Isaac, Guillermo Caruana, and Vicente Cuñat. Targeted product design. 2021.

Moshe A Barach, Joseph M Golden, and John J Horton. Steering in online markets: the role of platform incentives and credibility. *Management Science*, 2020.

Mathias Bärtl. Youtube channels, uploads and views: A statistical analysis of the past 10 years. *Convergence*, 24(1):16–32, 2018.

Matthew Blackwell, Stefano Iacus, Gary King, and Giuseppe Porro. cem: Coarsened exact matching in stata. *The Stata Journal*, 9(4):524–546, 2009.

David M Blei. Probabilistic topic models. *Communications of the ACM*, 55(4):77–84, 2012.

David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022, 2003.

Tommaso Bondi. Alone, together: Product discovery through consumer ratings. *Available at SSRN 3468433*, 2019.

Leif Brandes, David Godes, and Dina Mayzlin. What drives extremity bias in online reviews? theory and experimental evidence. *Theory and Experimental Evidence (September 6, 2019)*, 2019.

Heng Chen and Wing Suen. Competition for attention and news quality. Technical report, working paper, 2019.

Hsing Kenneth Cheng, Weiguo Fan, Peipei Guo, Hailiang Huang, and Liangfei Qiu. Can "gold medal" online sellers earn gold? the impact of reputation badges on sales. *Journal of Management Information Systems*, 37(4):1099–1127, 2020.

Rahul Deb, Matthew F Mitchell, and Mallesh Pai. (bad) reputation in relational contracting. 2020.

Daniel W Elfenbein, Ray Fisman, and Brian McManus. Charity as a substitute for reputation: Evidence from an online marketplace. *Review of Economic Studies*, 79(4):1441–1468, 2012.

Daniel W Elfenbein, Raymond Fisman, and Brian McManus. Market structure, reputation, and the value of quality certification. *American Economic Journal: Microeconomics*, 7(4):83–108, 2015.

Daniel Ershov and Matthew Mitchell. The effects of influencer advertising disclosure regulations: Evidence from instagram. In *Proceedings of the 21st ACM Conference on Economics and Computation*, pages 73–74, 2020.

Itay P Fainmesser and Andrea Galeotti. The market for influence. *Johns Hopkins Carey Business School Research*, pages 18–13, 2018.

Apostolos Filippas, John J Horton, and Joseph M Golden. Reputation inflation. Technical report, National Bureau of Economic Research, 2019.

Christopher Freeman and Luc Soete. *The economics of industrial innovation*. Psychology Press, 1997.

Spiros V Georgakopoulos, Sotiris K Tasoulis, Aristidis G Vrahatis, and Vassilis P Plagianakos. Convolutional neural networks for toxic comment classification. In *Proceedings of the 10th Hellenic Conference on Artificial Intelligence*, pages 1–6, 2018.

Timothy R Hannigan, Richard FJ Haans, Keyvan Vakili, Hovig Tchalian, Vern L Glaser, Milo Shaoqing Wang, Sarah Kaplan, and P Devereaux Jennings. Topic modeling in management research: Rendering new theory from textual data. *Academy of Management Annals*, 13(2):586–632, 2019.

Nan Hu, Jie Zhang, and Paul A Pavlou. Overcoming the j-shaped distribution of product reviews. *Communications of the ACM*, 52(10):144–147, 2009.

Influencer Marketing Hub. The state of influencer marketing 2021: benchmark report. *Influencer Marketing Hub, https://influencermarketinghub.com/influencer-marketing-benchmark-report-2021/, Accessed*, 2021.

Xiang Hui, Maryam Saeedi, Zeqian Shen, and Neel Sundaresan. Reputation and regulations: Evidence from ebay. *Management Science*, 62(12):3604–3616, 2016.

Xiang Hui, Zekun Liu, and Weiqing Zhang. Mitigating the cold-start problem in reputation systems: Evidence from a field experiment. *Available at SSRN*, 2020.

Stefano M Iacus, Gary King, and Giuseppe Porro. Causal inference without balance checking: Coarsened exact matching. *Political analysis*, pages 1–24, 2012.

Maarten C.W Jansen and Cole Williams. Influencing search. *working paper*, 2021.

Justin P Johnson and David P Myatt. On the simple economics of advertising, marketing, and product design. *American Economic Review*, 96(3):756–784, 2006.

Boyan Jovanovic. Job matching and the theory of turnover. *Journal of political economy*, 87(5, Part 1):972–990, 1979.

Zsolt Katona, Jonathan A Knee, and Miklos Sarvary. Agenda chasing and contests among news providers. *The RAND Journal of Economics*, 48(3):783–809, 2017.

Anna Kerkhof. Advertising and content differentiation: Evidence from youtube. *Available at SSRN 3733271*, 2020.

Gary King and Richard Alexander Nielsen. Why propensity scores should not be used for matching. 2019.

Tobias J Klein, Christian Lambertz, and Konrad O Stahl. Market transparency, adverse selection, and moral hazard. *Journal of Political Economy*, 124(6):1677–1713, 2016.

Tracy R Lewis and David EM Sappington. Supplying information to facilitate price discrimination. *International Economic Review*, pages 309–327, 1994.

Lingfang Ivy Li, Steven Tadelis, and Xiaolan Zhou. Buying reputation as a signal of quality: Evidence from an online marketplace. Technical report, National Bureau of Economic Research, 2016.

Michael Luca and Georgios Zervas. Fake it till you make it: Reputation, competition, and yelp review fraud. *Management Science*, 62(12):3412–3427, 2016.

James G March. Exploration and exploitation in organizational learning. *Organization science*, 2(1): 71–87, 1991.

Matthew Mitchell. Free (ad) vice: Internet influencers and disclosure regulation. *working paper*, 2020.

Amy Pei and Dina Mayzlin. Influencing the influencers. *Available at SSRN 3376904*, 2019.

Martin Reisenbichler and Thomas Reutterer. Topic modeling in marketing: recent advances and research opportunities. *Journal of Business Economics*, 89(3):327–356, 2019.

Paul R Rosenbaum and Donald B Rubin. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55, 1983.

Robert Seamans and Feng Zhu. Responses to entry in multi-sided markets: The impact of craigslist on local newspapers. *Management Science*, 60(2):476–493, 2014.

Carl Shapiro. Premiums for high quality products as returns to reputations. *The quarterly journal of economics*, 98(4):659–679, 1983.

Nachum Sicherman and Oded Galor. A theory of career mobility. *Journal of political economy*, 98 (1):169–192, 1990.

Monic Sun and Feng Zhu. Ad revenue and content commercialization: Evidence from blogs. *Management Science*, 59(10):2314–2331, 2013.

Shagun Tripathi and Harris Kyriakou. To share or not to share? assessing the impact of self-regulation on a peer-to-peer market. In *Academy of Management Proceedings*, volume 2020, page 19087. Academy of Management Briarcliff Manor, NY 10510, 2020.

Stefano Vaccari, Costis Maglaras, and Marco Scarsini. Social learning from online reviews with product choice. In *Proceedings of the 13th Workshop on Economics of Networks, Systems and Computation*, pages 1–1, 2018.

Yanshan Wang, Sunghwan Sohn, Sijia Liu, Feichen Shen, Liwei Wang, Elizabeth J Atkinson, Shreyasee Amin, and Hongfang Liu. A clinical text classification paradigm using weak supervision and deep representation. *BMC medical informatics and decision making*, 19(1):1, 2019.

Appendix A. Proofs

Proof of Lemma 1

Taking F.O.C with respect to x_t, we have:

$$\frac{\partial \Pi}{\partial x_t} = a_t[\phi(a_t + (1 - x_t) + \xi_t \Delta) - \phi(a_t + x_t + \xi_t \Delta)]$$

Apparently that when $x_t = \frac{1}{2}$, $\frac{\partial \Pi}{\partial x_t} = 0$, and the case of $x > \frac{1}{2}$ and $x < \frac{1}{2}$ are symmetric. When $a_t + \xi_t \Delta < \mu$, $\phi(a_t + (1 - x_t) + \xi_t \Delta) - \phi(a_t + x_t + \xi_t \Delta) > 0$ for all $x \in [0, \frac{1}{2})$. Therefore it is optimal to set the most broad design ($x_t = \frac{1}{2}$). When $a_t + \xi_t \Delta > \mu$, $\phi(a_t + (1 - x_t) + \xi_t \Delta) - \phi(a_t + x_t + \xi_t \Delta) < 0$ for all $x \in [0, \frac{1}{2})$, and it is optimal to set the most niche design $x_t = 0$ or $x_t = 1$.

Proof of Lemma 2

By the assumption of Log-concaveness, we have uniqueness of equilibrium, and $\frac{\partial a_t}{\partial \theta_t} > 0$ within both the most niche periods and the most broad periods.

For the connection between the niche periods and the broad periods, note that at $\bar{\theta}_t$, the influencer is indifferent between the niche and the broad design. This is the position where $a_t^*(N) = a_t^*(B) = a_t^*$. Therefore over the whole range of $\theta_t \in [0, 1]$, $\frac{\partial a_t}{\partial \theta_t} > 0$.

Proof of Lemma 3

We show by conduction. At $t = 0$, since the influencer is indifferent between type A and type B audiences, w.l.o.g let $x_0 \leq \frac{1}{2}$. Denote $\underline{u}_t^j$ as the lower bound of type $j \in \{A, B\}$ new subscribers' match value. $x_0 \leq \frac{1}{2}$ implies $\underline{u}_0^A \leq \underline{u}_0^B$. Assume that at time $t > 0$, for the history $H_t = \{(x_0, a_0), (x_1, a_1), \ldots, (x_{t-1}, a_{t-1})\}$, all the horizontal designs weakly lean to type A, that is, $x_s \leq \frac{1}{2}, \forall s \leq t - 1$. Then $\underline{u}_s^A \leq \underline{u}_s^B, \forall s \leq t - 1$. At period t, for any content design $\pi_t = (x_t, a_t)$ such that $x_t > \frac{1}{2}$, we have $\underline{u}_t^A(x_t, a_t) > \underline{u}_t^B(x_t, a_t)$. Then the total demand from the past subscribers is:

$$D(x_t, a_t) = \sum_{s=0}^{t-1} (1 - \gamma)^{(t-s)} [2 - \Phi(\max\{\underline{u}_s^A, \underline{u}_t^A(\pi_t)\}) - \Phi(\max\{\underline{u}_s^B, \underline{u}_t^B(\pi_t)\})]$$

Consider an alternative content design $\tilde{\pi}_t = (1 - x_t, a_t)$, by symmetry directly we have $\underline{u}_t^A(\tilde{x}_t, a_t) = \underline{u}_t^B(x_t, a_t)$ and $\underline{u}_t^B(\tilde{x}_t, a_t) = \underline{u}_t^A(x_t, a_t)$. And the total demand from the past subscribers is:

$$D(\tilde{x}_t, a_t) = \sum_{s=0}^{t-1}(1-\gamma)^{(t-s)}[2 - \Phi(\max\{\underline{u}_s^A, \underline{u}_t^A(\tilde{\pi}_t)\}) - \Phi(\max\{\underline{u}_s^B, \underline{u}_t^B(\tilde{\pi}_t)\})]$$

Let the s period subscribers' demand be $d(x_t, a_t) = 2 - \Phi(\max\{\underline{u}_s^A, \underline{u}_t^A(\pi_t)\}) - \Phi(\max\{\underline{u}_s^B, \underline{u}_t^B(\pi_t)\})$, and $d(\tilde{x}_t, a_t) = 2 - \Phi(\max\{\underline{u}_s^A, \underline{u}_t^A(\tilde{\pi}_t)\}) - \Phi(\max\{\underline{u}_s^B, \underline{u}_t^B(\tilde{\pi}_t)\})$. For any $s \leq t$, when $\min\{\underline{u}_t^A(\pi_t), \underline{u}_t^B(\pi_t)\} \geq \underline{u}_s^B$ or $\max\{\underline{u}_t^A(\pi_t), \underline{u}_t^B(\pi_t)\} \leq \underline{u}_s^A$, we have $d(x_t, a_t) = d(\tilde{x}_t, a_t)$. When $\underline{u}_s^A \leq \underline{u}_t^B(\pi_t) \leq \underline{u}_s^B \leq \underline{u}_t^B(\pi_t)$, we have $d(x_t, a_t) - d(\tilde{x}_t, a_t) = \Phi(\underline{u}_t^B(\pi_t)) - \Phi(\underline{u}_s^B) \leq 0$. When $\underline{u}_s^A \leq \underline{u}_t^B(\pi_t) \leq \underline{u}_t^B(\pi_t) \leq \underline{u}_s^B$, we have $d(x_t, a_t) - d(\tilde{x}_t, a_t) = \Phi(\underline{u}_t^B(\pi_t)) - \Phi(\underline{u}_t^A(\pi_t) \leq 0$. When $\underline{u}_t^B(\pi_t) \leq \underline{u}_s^A \leq \underline{u}_t^B(\pi_t) \leq \underline{u}_s^B$, we have $d(x_t, a_t) - d(\tilde{x}_t, a_t) = \Phi(\underline{u}_s^A) - \Phi(\underline{u}_t^A(\pi_t) \leq 0$. Lastly, when $\underline{u}_t^B(\pi_t) \leq \underline{u}_s^A \leq \underline{u}_s^B \leq \underline{u}_t^B(\pi_t)$, we have $d(x_t, a_t) - d(\tilde{x}_t, a_t) = \Phi(\underline{u}_s^A) - \Phi(\underline{u}_s^B) \leq 0$. Therefore, for any $s \leq t$, $d(x_t, a_t) \leq d(\tilde{x}_t, a_t)$ always holds. And the total demand $D(x_t, a_t) \leq D(\tilde{x}_t, a_t)$. Thus we show that $\tilde{\pi}_t = (1 - x_t, a_t)$ weakly dominates $\pi_t = (x_t, a_t)$ when $x_t \geq \frac{1}{2}$.

Proof of Lemma 4

Assume that in the $\gamma = 1$ section, the influencer switches to the broad design at period $\hat{t}$. Then when $\gamma < 1$, the influencer keep the most niche content design ($x_t = 0$) for at least $\hat{t}$ periods. To see that, denote the optimal design in the $\gamma = 1$ case (conditional on the realization is always good content) as $\{(a_0^\dagger, x_0^\dagger), (a_1^\dagger, x_1^\dagger)...\}$, and the optimal design in the $\gamma < 1$ case is $\{(a_0^*, x_0^*), (a_1^*, x_1^*)...\}$. Apparently $(a_0^*, x_0^*) = (a_0^\dagger, x_0^\dagger)$.

We first show that $x_t^* = 0, \forall t \leq \hat{t}$. Assume $x_1^* > x_1^\dagger = 0$. If a_1^* is such that $\underline{u}_1^B < \underline{u}_0^B$, then a profitable deviation is to remain a_1^* unchanged, deviate x_1^* to smaller direction while maintaining $\underline{u}_1^B \leq \underline{u}_0^B$. If a_1^* is such that $\underline{u}_1^B \geq \underline{u}_0^B$, it has to be $\underline{u}_1^A > \underline{u}_0^A$ as $x_1^* > x_1^\dagger = 0$. Thus a profitable deviation is again keeping a_1^* unchanged, while moving x_t to smaller direction. This proof naturally conveys to further periods until $\hat{t}$ by induction.

Knowing that $x_t^* = 0$, we next show that $\underline{u}_t^j$ monotonically decreases over time. Assume that the history at time t is $\underline{u}_0^j \geq \underline{u}_1^j \geq \cdots \geq \underline{u}_{t-1}^j$. At time t, the influencer chooses the optimal

$\underline{u}_t^A$[26] to maximize the period t static payoff. There are $t+1$ different sections that $\underline{u}_t^A$ can locate: $(\underline{u}_0^A, \infty), (\underline{u}_1^A, \underline{u}_0^A], (\underline{u}_2^A, \underline{u}_1^A], \dots, (\underline{u}_{t-1}^A, \underline{u}_{t-2}^A], (-\infty, \underline{u}_{t-1}^A]$ We show by contradiction that the only possible region that $\underline{u}_t^A$ locates is the last. If the optimal $\underline{u}_t^A$ is at region $(\underline{u}_s^A, \underline{u}_{s-1}^A]$ (where $0 \le s < t$, and $\underline{u}_{-1}^A = \infty$). The optimization problem for the period s influencer is:

$$\max_{\underline{u}_s^A}[2 - \Phi(\underline{u}_s^A) - \Phi(\underline{u}_s^A + 1) + \sum_{i=0}^{s-1}(1-\gamma)^{s-i}(2 - \Phi(\underline{u}_i^A) - \Phi(\underline{u}_i^A + 1))](\underline{u}_s^A - \xi_s \Delta)$$

The F.O.C (conditional on $\underline{u}_s^A \le \underline{u}_{s-1}^A$) results:

$$\underline{u}_s^A = \frac{2 + \sum_{i=0}^{s-1}(1-\gamma)^{s-i}(2 - \Phi(\underline{u}_i^A) - \Phi(\underline{u}_i^A + 1)) - \Phi(\underline{u}_s^A) - \Phi(\underline{u}_s^A + 1)}{\phi(\underline{u}_s^A) + \phi(\underline{u}_s^A + 1)} + \xi_s \Delta$$

At time t, if it is optimal for the influencer to set $\underline{u}_t^A \in (\underline{u}_s^A, \underline{u}_{s-1}^A]$, then the optimization problem within this region is:

$$\max_{\underline{u}_t^A}\left[2 - \Phi(\underline{u}_t^A) - \Phi(\underline{u}_t^A + 1) + \frac{\sum_{i=1}^{t-s}(1-\gamma)^i(2 - \Phi(\underline{u}_{s-1}^A) - \Phi(\underline{u}_{s-1}^A + 1)) + (1-\gamma)^{t-s}\sum_{i=0}^{s-1}(1-\gamma)^{s-i}(2 - \Phi(\underline{u}_i^A) - \Phi(\underline{u}_i^A + 1))}{\sum_{i=0}^{t-s}(1-\gamma)^i}\right](\underline{u}_t^A - \xi_t \Delta)$$

When $\underline{u}_0^j \ge \underline{u}_1^j \ge \cdots \ge \underline{u}_{t-1}^j$, it is easy to verify that:

$$\frac{\sum_{i=1}^{t-s}(1-\gamma)^i(2 - \Phi(\underline{u}_{s-1}^A) - \Phi(\underline{u}_{s-1}^A + 1)) + (1-\gamma)^{t-s}\sum_{i=0}^{s-1}(1-\gamma)^{s-i}(2 - \Phi(\underline{u}_i^A) - \Phi(\underline{u}_i^A + 1))}{\sum_{i=0}^{t-s}(1-\gamma)^i} < \sum_{i=0}^{s-1}(1-\gamma)^{s-i}(2 - \Phi(\underline{u}_i^A) - \Phi(\underline{u}_i^A + 1)) \tag{6}$$

And since the chance of producing a bad content at period t, ξ_t is decreasing over time for influencers, we have $\xi_t < \xi_s$. Both $\xi_t < \xi_s$ and inequality (6) drives $\underline{u}_t^A < \underline{u}_s^A$. We therefore show the existence of a contradiction.

Lastly, we show that while the influencer maintains the most niche content design, a_t monotonically increases. Assume that in all the history from $t = 0, 1, 2 \dots t$, a_t monotonically increases and $\underline{u}_t^A$ monotonically decreases. At time $t+1$, the influencer's optimization problem is:

$$a_{t+1} = \arg\max\left[2 - \Phi(a_{t+1} + \xi_{t+1}\Delta) - \Phi(a_{t+1} + \xi_{t+1}\Delta + 1) + \sum_{s=0}^{t}(1-\gamma)^{t+1-s}(2 - \Phi(\underline{u}_s^A) - \Phi(\underline{u}_s^A + 1))\right] a_{t+1}$$

[26] Note that when $x_t = 0$, $\underline{u}_t^B = \underline{u}_t^A + 1$ and $a_t = \underline{u}_t^A - x_t - \xi_t \Delta$.

The First Order Condition is:

$$a_{t+1}^* = \frac{2 + \sum_{s=0}^{t}(1-\gamma)^{t+1-s}(2 - \Phi(\underline{u}_s^A) - -\Phi(\underline{u}_s^A + 1)) - \Phi(a_{t+1} + \xi_{t+1}\Delta) - \Phi(a_{t+1} + \xi_{t+1}\Delta + 1)}{\phi(a_{t+1} + \xi_{t+1}\Delta) + \phi(a_{t+1} + \xi_{t+1}\Delta + 1)} - \xi_{t+1}\Delta$$

Compared to the F.O.C when the influencer is choosing the optimal advertising level at period t:

$$a_t^* = \frac{2 + \sum_{s=0}^{t-1}(1-\gamma)^{t-s}(2 - \Phi(\underline{u}_s^A) - \Phi(\underline{u}_s^A + 1)) - \Phi(a_t + \xi_t\Delta) - \Phi(a_t + \xi_t\Delta + 1)}{\phi(a_t + \xi_t\Delta) + \phi(a_t + \xi_t\Delta + 1)} - \xi_t\Delta$$

Since $\underline{u}_t^A$ is monotonically decreasing in t, we have $\sum_{s=0}^{t-1}(1-\gamma)^{t-s}(1-\Phi(\underline{u}_s^A)) < \sum_{s=0}^{t}(1-\gamma)^{t+1-s}(1-\Phi(\underline{u}_s^A))$. For a survival influencer, since he always produce good content, $\xi_{t+1} < \xi_t$ also holds. Therefore $a_{t+1}^* > a_t^*$.

Proof of Proposition 1

Let the first period the influencer switches to non-niche design is at time $\hat{t}$. We divide the proof into two cases. The first case is that $\hat{t} = 0$, where the influencer is optimal to start broad from $t = 0$. The second case is when $\hat{t} > 0$. The influencer begins with most niche design ($t = 0$) for $\hat{t}$ periods, and switching to broad design at $\hat{t}$.

Case 1: $\hat{t} = 0$**.** This case happens when the influencer has high enough initial belief θ_0 to be H type. Lemma 1 guarantees that the influencer sets the most broad design, $x_0^* = \frac{1}{2}$ at $t = 0$. We first show that the product design is always most broad ($x_t = \frac{1}{2}$). Since the product design at $t = 0$ is most broad, the subscribers from $t = 0$ are balanced between two types, and the lowest match-value subscriber of both types is $\underline{u}_0^A = \underline{u}_0^B$. At $t = 1$, assume that $x_1 < \frac{1}{2}$. A profitable deviation for the influencer is to keep a_1 unchanged and increase x_1 such that $\underline{u}_1^A \leq \underline{u}_1^B$. This proof conveys to further periods by induction.

Next we show that a_t increases and $\underline{u}_t^A = \underline{u}_t^B$ decreases over time. Knowing that $x_t = \frac{1}{2}, \forall t$, assume that at time t, $a_0^* < a_1^* < a_2^* < \cdots < a_t^*$, and $\underline{u}_0^A \geq \underline{u}_1^A \geq \cdots \geq \underline{u}_t^A$. We aim to show that at $t+1$, $a_{t+1}^* > a_t^*$. For the exact same argument of Lemma 4, we know that $\underline{u}_{t+1}^A \leq \underline{u}_t^A$. And the influencer's optimization problem is:

$$a_{t+1} = \arg\max \left[1 - \Phi(a_{t+1} + \xi_{t+1}\Delta + \frac{1}{2}) + \sum_{s=0}^{t}(1-\gamma)^{t+1-s}(1 - \Phi(\underline{u}_s^A))\right] a_{t+1} \tag{7}$$

The First Order Condition is:

$$a_{t+1}^* = \frac{1 + \sum_{s=0}^{t}(1-\gamma)^{t+1-s}(1 - \Phi(\underline{u}_s^A)) - \Phi(a_{t+1} + \xi_{t+1}\Delta + \frac{1}{2})}{\phi(a_{t+1} + \xi_{t+1}\Delta + \frac{1}{2})} - \xi_{t+1}\Delta \tag{8}$$

Compared to the F.O.C when the influencer is choosing the optimal advertising level at period t:

$$a_t^* = \frac{1 + \sum_{s=0}^{t-1}(1-\gamma)^{t-s}(1 - \Phi(\underline{u}_s^A)) - \Phi(a_t + \xi_t\Delta + \frac{1}{2})}{\phi(a_t + \xi_t\Delta + \frac{1}{2})} - \xi_t\Delta \tag{9}$$

Since $\underline{u}_t^A$ is monotonically decreasing in t, we have $\sum_{s=0}^{t-1}(1-\gamma)^{t-s}(1-\Phi(\underline{u}_s^A)) < \sum_{s=0}^{t}(1-\gamma)^{t+1-s}(1-\Phi(\underline{u}_s^A))$. For a survival influencer, since he always produce good content, $\xi_{t+1} < \xi_t$ also holds. Therefore $a_{t+1}^* > a_t^*$.

Case 2: $\hat{t} > 0$. For the exact same argument of Case 1, we know that during the beginning $\hat{t}$ periods, $a_0^* < a_1^* < \cdots < a_{\hat{t}-1}^*$, and $\underline{u}_0^A \geq \underline{u}_1^A \geq \cdots \geq \underline{u}_{\hat{t}-1}^A$. Also since $x_0^* = x_1^* = \cdots = x_{\hat{t}-1}^* = 0$, we have $\underline{u}_t^B = \underline{u}_t^A + 1, \forall t < \hat{t}$. Again, our interest is to examine the dynamics of the content design, (x_t, a_t) and the lower boundary of the audiences who consume the content at each period, $(\underline{u}_t^A, \underline{u}_t^B)$.

We first show that no past audiences will be abandoned at any $t \geq \hat{t}$, that is, $\forall t \geq \hat{t}$, $u_t^i \leq u_{t-1}^i, i = A, B$. To begin with, note that $\underline{u}_{\hat{t}-1}^A + \underline{u}_{\hat{t}-1}^B \geq 2\mu$. This is because the match value distribution $\phi(\cdot)$ is symmetric to μ and log-concave. If $\underline{u}_{\hat{t}-1}^A + \underline{u}_{\hat{t}-1}^B < 2\mu$, there must exists a time $\bar{t} < \hat{t}$, such that $\underline{u}_{\bar{t}}^A < \mu - \frac{1}{2} \leq \underline{u}_{\bar{t}-1}^A$. Then a profitable deviation at time $\bar{t}$ is to keep $a_{\bar{t}}$ unchanged and increasing $x_{\bar{t}}$, such that $\underline{u}_{\bar{t}}^A = \mu - \frac{1}{2}$. There are two cases that some audiences are abandoned and we show both cannot hold. First, if at time $\hat{t}$, $\underline{u}_{\hat{t}}^i \geq \underline{u}_{\hat{t}-1}^i$ for both $i = A, B$ with at least one strict inequality, then a profitable deviation is to first keep $a_{\hat{t}}$ unchanged and decrease x_t to 0. And secondly decrease $a_{\hat{t}}$ such that $\underline{u}_{\hat{t}}^i = \underline{u}_{\hat{t}-1}^i$. The second step is directly from Lemma 4. Second, if at time $\hat{t}$, $\underline{u}_{\hat{t}}^A > \underline{u}_{\hat{t}-1}^A$ and $\underline{u}_{\hat{t}}^B \leq \underline{u}_{\hat{t}-1}^B$. By assumption 1, we know that

$(1+\gamma)\phi(\underline{u}_{\hat{t}}^A) > (1+\gamma)\phi(\mu - \frac{1}{2}) \geq \phi(\mu) \geq \phi(\underline{u}_{\hat{t}}^B)$. Therefore a profitable deviation is to keep $a_{\hat{t}}$ unchanged, and decrease $x_{\hat{t}}$ till $\underline{u}_{\hat{t}}^A = \underline{u}_{\hat{t}-1}^A$.

Knowing that the influencer would not abandon past audiences, the economic trade-offs between x_t and a_t is purely on the new audiences and we next show the changes in the x_t and a_t. First, there can only be two cases: $\underline{u}_t^A = \underline{u}_{t-1}^A$ and $x_{t-1} \leq x_t \leq \frac{1}{2}$, or $\underline{u}_t^A > \underline{u}_{t-1}^A$ and $x_t = \frac{1}{2}$. To see that, note that at $t \geq \hat{t}$, when we only consider the newly arrived audiences, by Lemma 1 the influencer optimally chooses the broad design. Therefore if $\underline{u}_t^A > \underline{u}_{t-1}^A$ and $x_t < \frac{1}{2}$, it would be profitable for the influencer to keep a_t unchanged and increase x_t, until $x_t = \frac{1}{2}$ or $\underline{u}_t^A = \underline{u}_{t-1}^A$. Thus in this case x_t monotonically increases.

Next, we show that $a_t \geq a_{t-1}$. First we consider the case that $\underline{u}_t^A = \underline{u}_{t-1}^A$. A higher a_t induces a lower $\underline{u}_t^B$. More specifically, when $\underline{u}_t^A = \underline{u}_{t-1}^A$, it implies that $a_t + x_t + \xi_t\Delta = a_{t-1} + x_{t-1} + \xi_{t-1}\Delta$. Therefore $x_t = (a_{t-1} - a_t) + (\xi_{t-1} - \xi_t)\Delta + x_{t-1}$, and $\underline{u}_t^B = 1 + 2a_t + 2\xi_t\Delta - \underline{u}_{t-1}^A$ The influencer's optimization problem becomes:

$$\max_{a_t} \left[2 - \Phi(\underline{u}_{t-1}^A) - \Phi(1 + 2a_t + 2\xi_t\Delta - \underline{u}_{t-1}^A) + \sum_{i=0}^{t-1} (1-\gamma)^{t-i} (2 - \Phi(\underline{u}_i^A) - \Phi(\underline{u}_i^B)) \right] a_t$$

The first order condition that determines a_t (if the solution is interior) is:

$$a_t = \frac{2 - \Phi(\underline{u}_{t-1}^A) + \sum_{i=0}^{t-1} (1-\gamma)^{t-i} (2 - \Phi(\underline{u}_i^A) - \Phi(\underline{u}_i^B)) - \Phi(1 + 2a_t + 2\xi_t\Delta - \underline{u}_{t-1}^A)}{2\phi(1 + 2a_t + 2\xi_t\Delta - \underline{u}_{t-1}^A)}$$

Since the public belief of survival influencers monotonically grows, $\xi_t < \xi_{t-1}$. when the difference between ξ_t and ξ_{t-1} is small, the optimal a_t leads to a corner solution that $\underline{u}_t^B = \underline{u}_{t-1}^B$ and $a_t > a_{t-1}$. When ξ_t is getting smaller so that there is interior solution, the F.O.C leads to a even greater a_t. Lastly we consider the case where $\xi_{t-1} - \xi_t$ is large enough such that the optimal $x_t = \frac{1}{2}$. There exists an $\underline{\xi}_t$, such that at $\underline{\xi}_t$, the optimal content design is such that $\underline{u}_t^A = \underline{u}_t^B = \underline{u}_{t-1}^A$. If $\xi_t < \underline{\xi}_t$, the optimal design would be such that $\underline{u}_t^A = \underline{u}_t^B < \underline{u}_{t-1}^A$. It is easy to see that when ξ_t decreases, it monotonically increase a_t. Therefore for all cases we show that $a_t > a_{t-1}$.

Proof of Proposition 2

We compare two influencers at t and t' respectively, and $t' > t$. The history of optimal policies at time t' is $H_{t'} = \{(x_0^*, a_0^*), (x_1^*, a_1^*), \ldots, (x_{t'-1}^*, a_{t'-1}^*)\}$, which associates with the lower bar of each type subscribers, $\underline{u}_0^j \geq \underline{u}_1^j \geq \cdots \geq \underline{u}_{t'-1}^j$, $j = A, B$. We aim to show that given the history up to $t-1$ and $t'-1$ respectively, the older influencer would allocate more share of content to the new event. That is, $s_t^N \leq s_{t'}^N$, $\forall t < t'$.

First, we show that $\forall t < t'$, under the optimal strategy $\underline{u}_t^j(a_t^*, x_t^*, s_t^{N*}) \geq \underline{u}_{t'}^j(a_{t'}^*, x_{t'}^*, s_{t'}^{N*})$. Note that given $(\underline{u}_t^A, \underline{u}_t^B)$, the influencer chooses the optimal s_t^N to maximize a_t. Let $\bar{u}_t = \frac{1}{2}(\underline{u}_t^A + \underline{u}_t^B)$, we have:

$$a_t = \bar{u}_t(1 - s_t^N) + u_N(s_t^N) - \xi_t \Delta$$

Taking F.O.C with respect to s_t^N, $\frac{\partial a_t}{\partial s_t^N} = -\bar{u}_t + u_N'(s_t^N) = 0$, we have $s_t^{N*} = u_N'^{-1}(\bar{u}_t)$. Therefore we can write the optimal a_t given $(\underline{u}_t^A, \underline{u}_t^B, \xi_t)$ as follows:

$$a_t^*(\underline{u}_t^A, \underline{u}_t^B, \xi_t) = \frac{1}{2}(\underline{u}_t^A + \underline{u}_t^B)(1 - u_N'^{-1}(\bar{u}_t)) + u_N(u_N'^{-1}(\bar{u}_t)) - \xi_t \Delta$$

Thus the game goes back to the similar problem as in Proposition 1: given the history of past subscribers up to time t, the influencer chooses optimal $(\underline{u}_t^A, \underline{u}_t^B)$ to maximize static payoff. More specifically, the influencer faces the following optimization problem:

$$\max_{\underline{u}_t^A, \underline{u}_t^B} \left[2 - \Phi(\underline{u}_t^A) - \Phi(\underline{u}_t^B) + \sum_{i=0}^{t-1}(1-\gamma)^{t-i}(2 - \Phi(\underline{u}_i^A) - \Phi(\underline{u}_i^B))\right] a_t^*(\underline{u}_t^A, \underline{u}_t^B, \xi_t)$$

The first order conditions w.r.t $(\underline{u}_t^A, \underline{u}_t^B)$ gives:

$$\frac{\partial a_t^*(\underline{u}_t^{A*}, \underline{u}_t^{B*}, \xi_t)}{\partial \underline{u}_t^j} \left[2 - \Phi(\underline{u}_t^{A*}) - \Phi(\underline{u}_t^{B*}) + \sum_{i=0}^{t-1}(1-\gamma)^{t-i}(2 - \Phi(\underline{u}_i^A) - \Phi(\underline{u}_i^B))\right] = \phi(\underline{u}_t^j) a_t^*(\underline{u}_t^{A*}, \underline{u}_t^{B*}, \xi_t) \tag{10}$$

for $j = A, B$. The LHS of (10) is the marginal gain of increase $\underline{u}_t^j$ and the RHS of (10) is the marginal loss. Assume that at $t' > t$, the influencer applies the same strategy: $(\underline{u}_{t'}^A, \underline{u}_{t'}^B) = (\underline{u}_t^{A*}, \underline{u}_t^{B*})$. The marginal loss increases (as $\xi_{t'} < \xi_t$), and the marginal gain decreases. Therefore the influencer has strict incentive to reduce both lower bounds such that $(\underline{u}_{t'}^{A*}, \underline{u}_{t'}^{B*}) < (\underline{u}_t^{A*}, \underline{u}_t^{B*})$

Since $u_N(\cdot)$ is increasing and concave, $u_N'^{-1}(\cdot)$ is decreasing. Therefore by $\bar{u}_t \geq \bar{u}_{t'}$ we have $s_t^{N*} \leq s_{t'}^{N*}$. That is, a young influencer allocates more contents to the new event than an old influencer.

Proof of Proposition 3

There are three parts in this proposition: unbiased quality Q_t decreasing, selection bias Δ_t decreasing, and the decreasing difference for the rating bias. Note that the first part is straightforward from $Q_t = \mu - a_t - \frac{1}{2} + \epsilon_t$ and Proposition 1. We only need to examine the dynamics of $\Delta_t(\epsilon_t)$. Δ_t can be written as:

$$\begin{aligned}\Delta_t(\epsilon_t) = \mathcal{R}_{it} - Q_t \quad &= \frac{(1+\alpha G(\epsilon_t))\left[\int_{\underline{u}_t^A}^{\infty}[u_i+\frac{1}{2}-x_t]dF_t^A(u_i)+\int_{\underline{u}_t^B}^{\infty}[u_i+\frac{1}{2}-(1-x_t)]dF_t^B(u_i)\right]}{\left[\int_{\underline{u}_t^A}^{\infty}dF_t^A(u_i)+\int_{\underline{u}_t^B}^{\infty}dF_t^B(u_i)\right](1+G(\epsilon_t))} \\ &= \frac{1+\alpha G(\epsilon_t)}{1+G(\epsilon_t)}\left[\frac{1}{2}+\frac{\int_{\underline{u}_t^A}^{\infty}u_i dF_t^A(u_i)+\int_{\underline{u}_t^B}^{\infty}u_i dF_t^B(u_i)}{\int_{\underline{u}_t^A}^{\infty}dF_t^A(u_i)+\int_{\underline{u}_t^B}^{\infty}dF_t^B(u_i)}+\frac{-x_t\int_{\underline{u}_t^A}^{\infty}dF_t^A(u_i)-(1-x_t)\int_{\underline{u}_t^B}^{\infty}dF_t^B(u_i)}{\int_{\underline{u}_t^A}^{\infty}dF_t^A(u_i)+\int_{\underline{u}_t^B}^{\infty}dF_t^B(u_i)}\right]\end{aligned} \tag{11}$$

The first term of (11) remains the same over time for given ϵ. The second term is nothing but the average match value of subscribers at time t. The distribution of subscribers, $F_t^i(u), i = A, B$ is the depreciated summation of all the past audiences. More specifically:

$$f_t^i(u) = \sum_{s=0}^{t}(1-\gamma)^{t-s}\mathbb{1}(u \geq u_s^i)\phi(u)$$

for $i = A, B$ and $\int f_t^i(u)du = F_t^i(u)$. From the proof of the Proposition 1, both $\underline{u}_t^A$ and $\underline{u}_t^B$ monotonically decreases over time, and this leads to a decreasing second term of (11). Note that from the proof of Proposition 1, we have $\underline{u}_t^A \leq \underline{u}_t^B$ for all t. Thus the accumulation of type A subscribers, $\int_{\underline{u}_t^A}^{\infty} dF_t^A(u_i)$, is greater than that of type B for all t. Since x_t monotonically increases from 0 to $\frac{1}{2}$ over time, the third term of (11) also monotonically decreasing over time. Thus we prove that $\frac{\partial\Delta_t(\epsilon)}{\partial t} < 0, \forall\epsilon$.

Lastly we show the decreasing difference. We define $\Delta_t(\epsilon) = \frac{\tilde{\Delta}_t}{1+G(\epsilon)}$. It is easy to see from (11) that $\tilde{\Delta}_t$ is decreasing over time and does not contain ϵ. For any $\epsilon < \epsilon'$, we have:

$$\Delta_t(\epsilon) - \Delta_t(\epsilon') = \tilde{\Delta}_t \frac{G(\epsilon') - G(\epsilon)}{(1 + G(\epsilon))(1 + G(\epsilon'))}$$

Since $0 < \frac{G(\epsilon') - G(\epsilon)}{(1+G(\epsilon))(1+G(\epsilon'))} < 1$ and for all $t < t'$, $\tilde{\Delta}_t \geq \tilde{\Delta}_{t'}$, we have $\Delta_t(\epsilon) - \Delta_t(\epsilon') \geq \Delta_{t'}(\epsilon) - \Delta_{t'}(\epsilon')$.

Appendix B. Figures and Tables

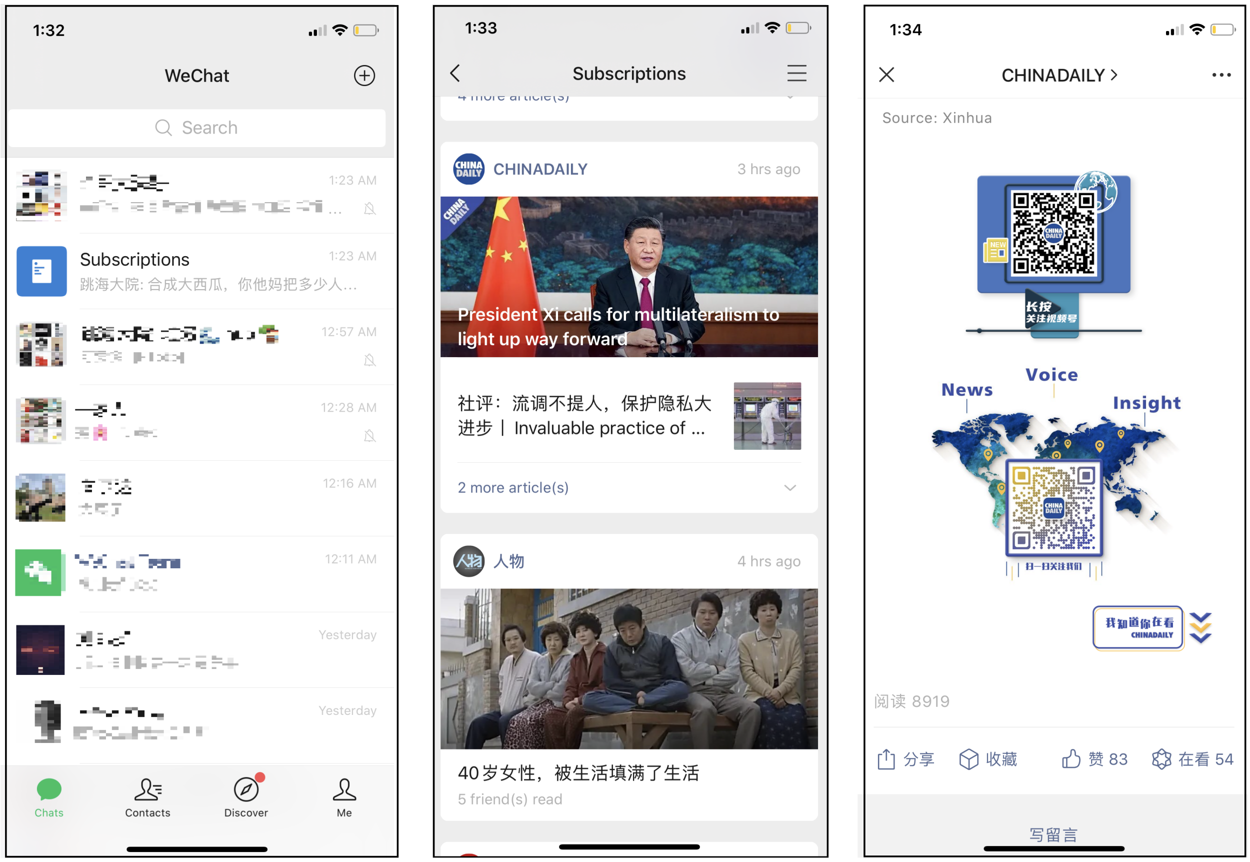

Figure I: Wechat Official Accounts Interface. The left panel is main interface of Wechat and irrelevant private chats are pixelated. Clicking "Subscriptions" navigates users to the official accounts that they have subscribed (middle panel). The right panel shows the number of views (8919) and likes (83) at the bottom of each article.

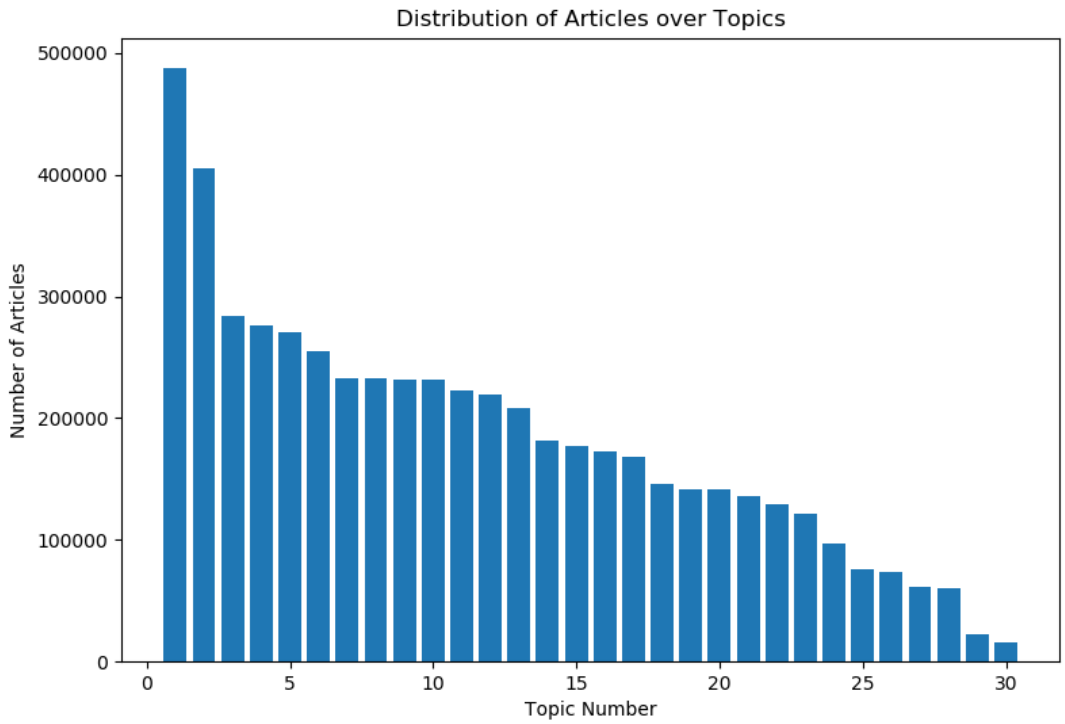

Figure II: Distribution of Articles over the Main Topics (sorted by number of articles)

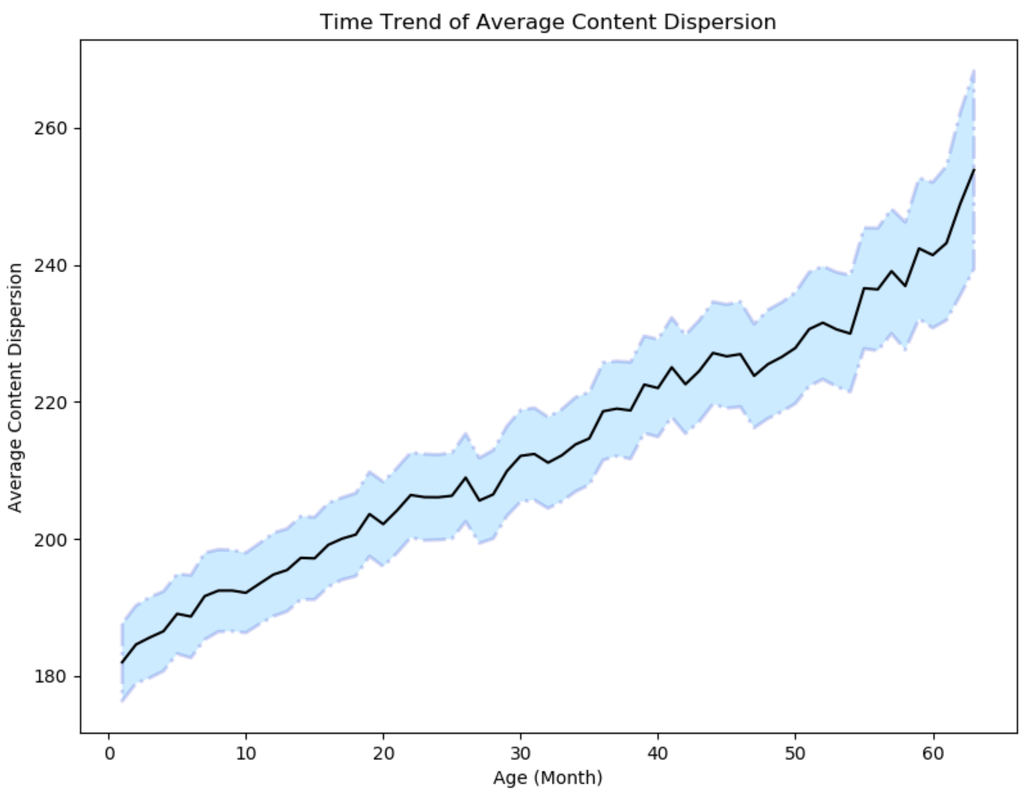

Figure III: Time Trend of Average Dispersion in Content Topics with 95% Confidence Interval

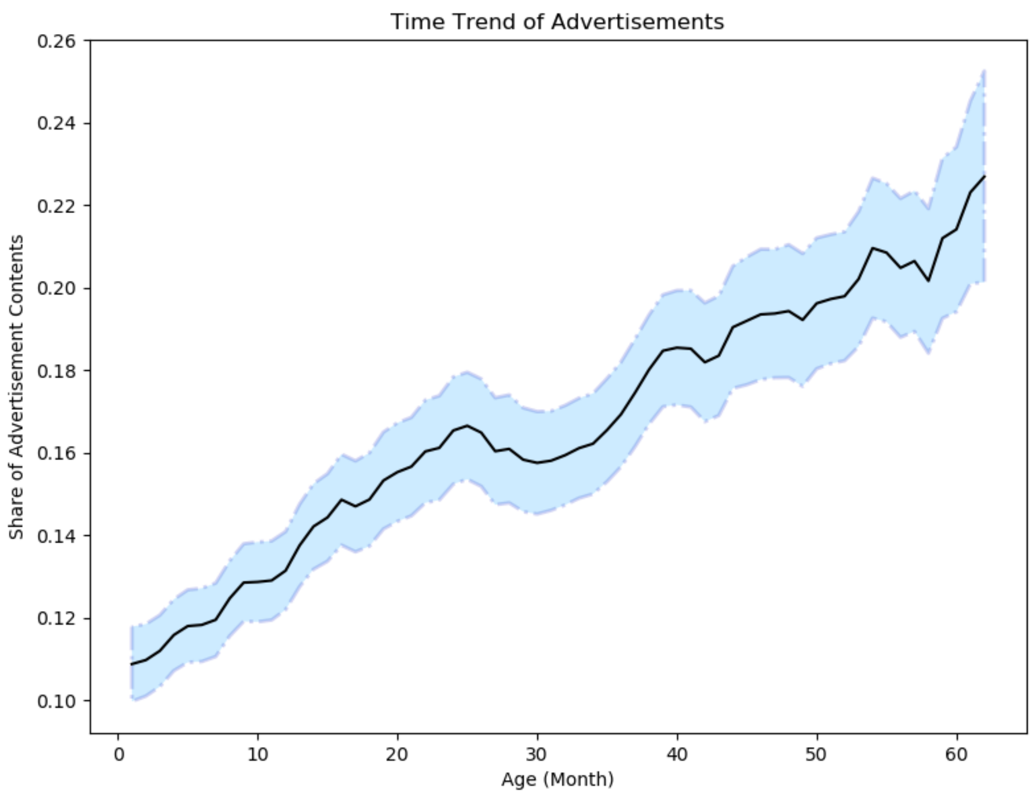

Figure IV: Time Trend of Average Advertisement Share with 95% Confidence Interval

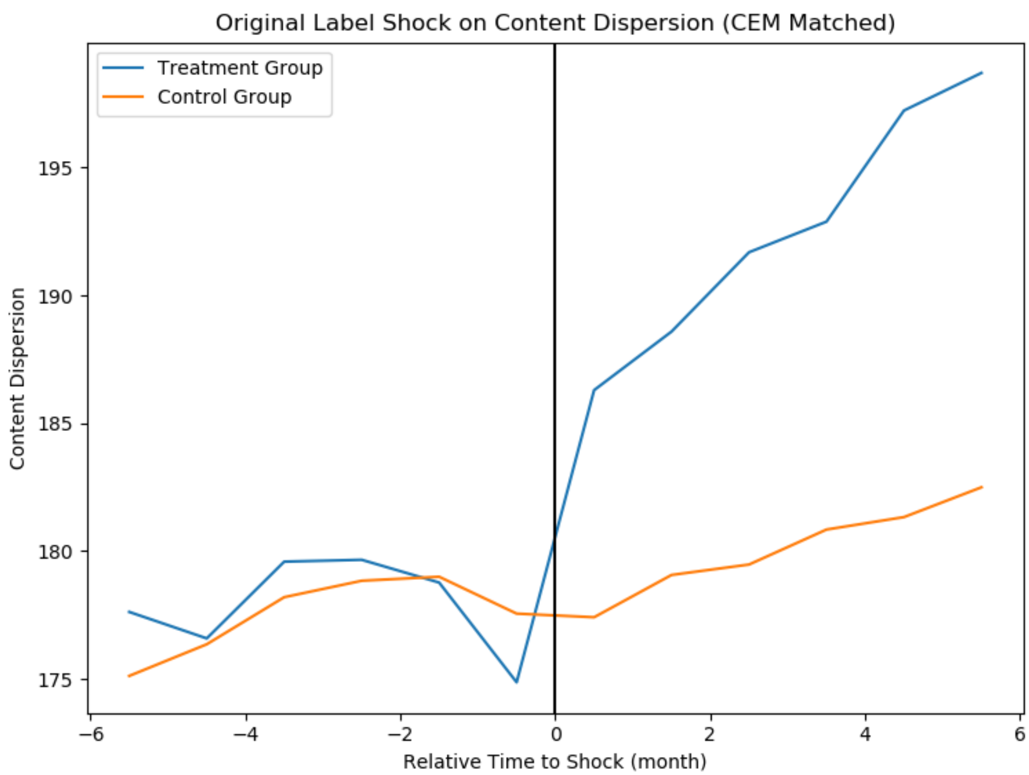

Figure V: Original tag shock on content dispersion (measured by average Euclidean distance) with control group identified by CEM method

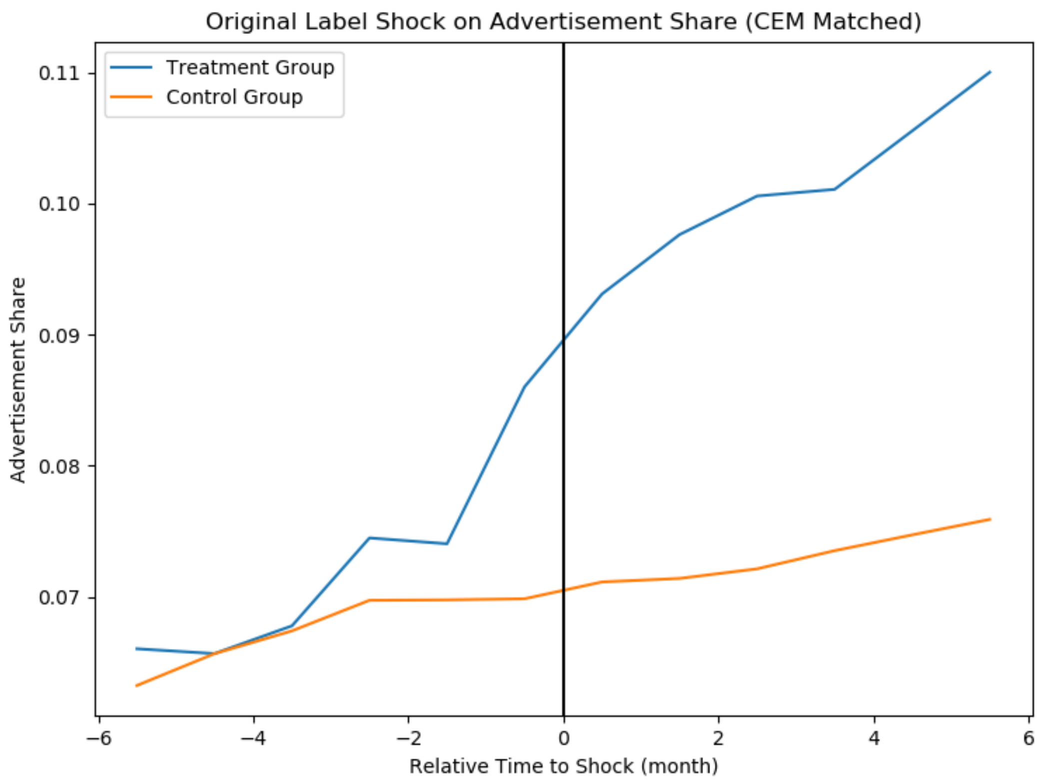

Figure VI: Original tag shock on share of advertising posts with control group identified by CEM method

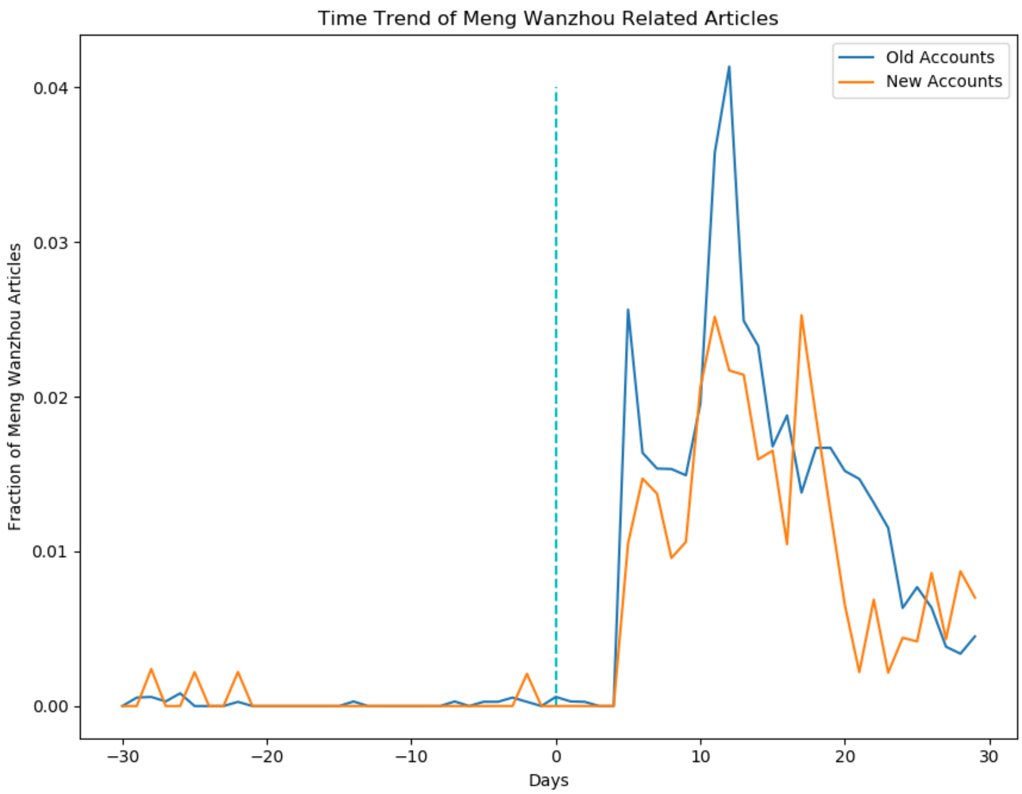

Figure VII: Time Trend of Meng Wanzhou Articles between Old and New Accounts

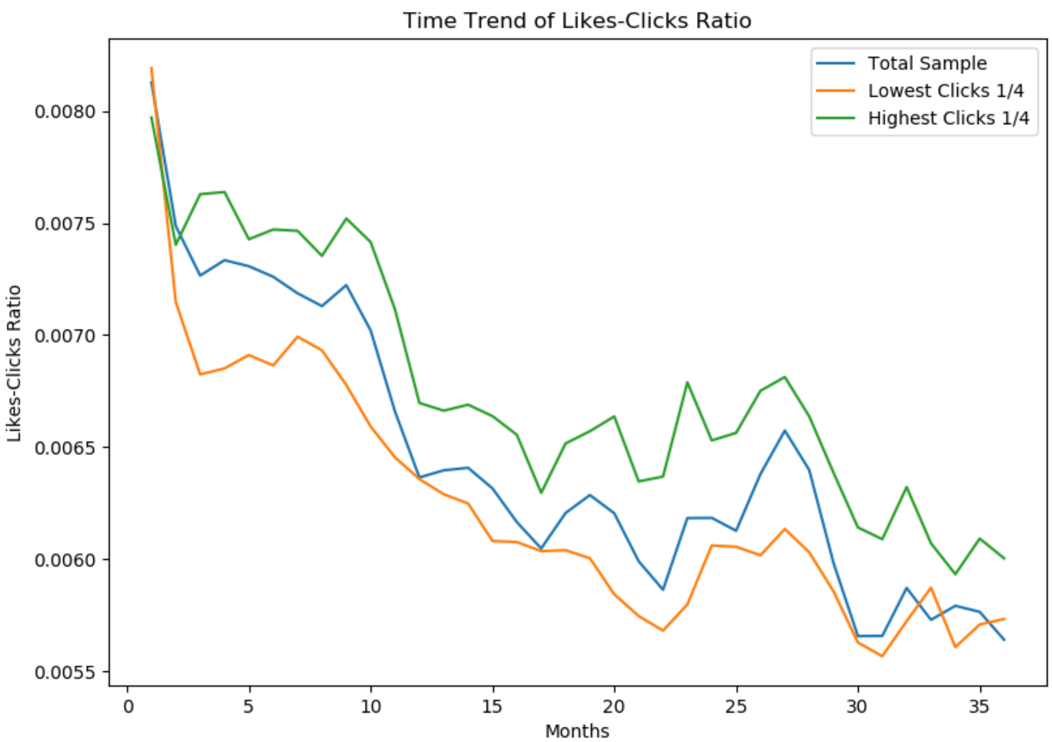

Figure VIII: Average Likes-Clicks Trend of Different Quantiles in Clicks

Table I: Summary Statistics of WeChat Data

Variable	Mean	Std Dev	Min	Max
# words	926.10	1246.70	0	29226
# likes	368.43	1276.26	0	100001
# clicks	34068.45	32049.52	0	100001
order number	2.40	2.04	0	7
original flag	0.15	0.36	0	1
# days since Sep14	1163.77	503.47	0	1919
pub year	2017.34	1.38	2014	2019
pub month	6.71	3.39	1	12
pub day	15.83	8.79	1	31
pub hour	15.09	6.44	0	23

Table II: High Weight Keywords of Each Topic

Topics	Keywords (translations are below) and Weights
Topic 1	0.017*″女人″ + 0.015*″爱″ + 0.015*″男人″ + 0.009*″喜欢″ + 0.008*″婚姻″ + 0.007*″爱情″ + 0.006*″结婚″ + 0.006*″朋友″ women, love, men, like, marriage, affection, get married, friends
Topic 2	0.062*″孩子″ + 0.014*″父母″ + 0.013*″妈妈″ + 0.008*″人生″ + 0.007*″家长″ + 0.007*″女儿″ + 0.006*″儿子″ + 0.006*″爱″ children, father and mother, mother, life, parents, daughter, son, love
Topic 3	0.008*″网友″ + 0.006*″明星″ + 0.005*″粉丝″ + 0.005*″拍″ + 0.004*″节目″ + 0.004*″真的″ + 0.004*″脸″ + 0.003*″娱乐圈″ netizen, celebrity, fans, shoot, show, true, face, entertainment
Topic 4	0.008*″男子″ + 0.007*″发生″ + 0.007*″视频″ + 0.006*″分″ + 0.006*″司机″ + 0.006*″警察″ + 0.005*″警方″ + 0.005*″网友″ man, happen, video, points, driver, police officer, police force, netizen
Topic 5	0.019*″中国″ + 0.009*″市场″ + 0.008*″美国″ + 0.008*″公司″ + 0.007*″经济″ + 0.007*″企业″ + 0.004*″行业″ + 0.004*″数据″ China, market, United States, company, economy enterprise, industry, data
Topic 6	0.018*″吃 + 0.007*″身体″ + 0.006*″疾病″ + 0.005*″食物″ + 0.005*″作用″ + 0.005*″人体″ + 0.004*″养生″ + 0.004*″治疗″ eat, body, disease, food, effect, human body, health, therapy
Topic 7	0.008*″书″ + 0.008*″时间″ + 0.007*″人生″ + 0.006*″读书″ + 0.006*″能力″ + 0.005*″读″ + 0.005*″越″ + 0.005*″世界″ book, time, life, read book, capability, read, more, world
Topic 8	0.016*″手机″ + 0.007*″车型″ + 0.007*″汽车″ + 0.007*″苹果″ + 0.006*″iPhone″ + 0.006*″SUV″ + 0.005*″车″ + 0.005*″发动机″ cellphone, car type, automobile, Apple, iPhone, SUV, cars, engine
Topic 9	0.009*″听 + 0.008*″音乐″ + 0.006*″经典″ + 0.005*″世界″ + 0.004*″爱″ + 0.004*″唱″ + 0.004*″人生″ + 0.004*″歌″ listen, music, classical, world, love, sing, life, song
Topic 10	0.008*″皮肤″ + 0.007*″买″ + 0.006*″效果″ + 0.006*″产品″ + 0.006*″肌肤″ + 0.005*″购买″ + 0.005*″面膜″ + 0.004*″成分″ skin, buy, effect, product, skin, purchase, facial mask, ingredient
Topic 11	0.008*″心″ + 0.006*″人生″ + 0.005*″字″ + 0.004*″文化″ + 0.003*″智慧″ + 0.003*″事″ + 0.003*″佛″ + 0.003*″菩萨″ heart, life, character, culture, wisdom, objects, buddha, bodhisattva
Topic 12	0.007*″深圳″ + 0.006*″路″ + 0.006*″北京″ + 0.005*″城市″ + 0.005*″时间″ + 0.004*″上海″ + 0.004*″地铁″ + 0.003*″站″ Shenzhen, road, Beijing, city, time, Shanghai, subway, station
Topic 13	0.014*″电影″ + 0.006*″故事″ + 0.006*″美国″ + 0.006*″中国″ + 0.005*″世界″ + 0.004*″导演″ + 0.003*″游戏″ + 0.003*″作品″ movie, story, United States, China, world, director, game, works
Topic 14	0.024*″穿″ + 0.011*″搭配″ + 0.011*″搭″ + 0.008*″时尚″ + 0.007*″衣服″ + 0.006*″好看″ + 0.005*″女人″ + 0.005*″时髦″ wear, collocation, pair, fashion, clothes, good looking, women, fashionable
Topic 15	0.007*″走″ + 0.005*″笑″ + 0.005*″听″ + 0.003*″问″ + 0.003*″看着″ + 0.003*″歌″ + 0.003*″回来″ + 0.003*″东西″ walk, laugh, listen, ask, look at, songs, come back, items
Topic 16	0.014*″学习″ + 0.010*″老师″ + 0.010*″学生″ + 0.007*″课程″ + 0.007*″中国″ + 0.006*″大学″ + 0.005*″数学″ + 0.005*″学校″ study, teacher, student, lesson, China, university, mathematics, school
Topic 17	0.039*″中国″ + 0.013*″美国″ + 0.009*″日本″ + 0.007*″国家″ + 0.006*″历史″ + 0.004*″世界″ + 0.004*″印度″ + 0.004*″俄罗斯″ China, United States, Japan, country, history, world, India, Russia
Topic 18	0.026*″吃 + 0.006*″煮″ + 0.006*″肉″ + 0.005*″锅″ + 0.005*″水″ + 0.005*″茶″ + 0.004*″好吃″ + 0.004*″鸡蛋″ eat, cook, meat, pot, water, tea, tasty, egg
Topic 19	0.016*″吃″ + 0.012*″店″ + 0.006*″广州″ + 0.006*″地址″ + 0.006*″美食″ + 0.005*″餐厅″ + 0.004*″酒店″ + 0.004*″味道″ eat, store, Guangzhou, address, gourmet, restaurants, hotel, taste
Topic 20	0.019*″钱 + 0.010*″医院″ + 0.008*″买″ + 0.008*″医生″ + 0.004*″理财″ + 0.004*″信息″ + 0.004*″医疗″ + 0.004*″花″ money, hospital, buy, doctors, financial management, information, medical treatment, spend
Topic 21	0.008*″报名″ + 0.008*″福利″ + 0.006*″送″ + 0.006*″时间″ + 0.005*″玩″ + 0.005*″参与″ + 0.005*″朋友圈″ + 0.005*″现场″ sign up, benefits, give away, time, play, participant, Wechat moment, on site
Topic 22	0.021*″公司″ + 0.010*″员工″ + 0.009*″企业″ + 0.008*″老板″ + 0.007*″招聘″ + 0.007*″创业″ + 0.006*″华为″ + 0.006*″团队″ company, employee, enterprise, boss, recruitment, start a business, Huawei, team
Topic 23	0.020*″成″ + 0.020*″朋友″ + 0.018*″谢谢″ + 0.013*″送给″ + 0.012*″星座″ + 0.011*″送″ + 0.010*″祝福″ + 0.009*″收到″ become, friend, thanks, give present, constellation, give, blessing, receive
Topic 24	0.031*″宝宝″ + 0.012*″瑜伽″ + 0.01*″动作″ + 0.009*″腿″ + 0.009*″身体″ + 0.008*″健身″ + 0.008*″妈妈″ + 0.007*″肌肉″ baby, yoga, action, leg, body, bodybuilding, mother, muscle
Topic 25	0.016*″粉丝″ + 0.016*″韩国″ + 0.008*″驴″ + 0.006*″散打″ + 0.006*″仙洋″ + 0.006*″天道″ + 0.006*″漫画″ + 0.005*″回归″ fans, Korea, donkey, sanshou, Xianyang (an internet celebrity), Tiandao (an internet celebrity), comics, come back
Topic 26	0.018*″哥″ + 0.009*″内衣″ + 0.008*″酒″ + 0.007*″阿哲″ + 0.007*″YY″ + 0.006*″粉丝″ + 0.006*″官方″ + 0.06*女人 Brother, underwear, liquor, Azhe (name of a streamer), YY (a live stream platform), fans, official, women
Topic 27	0.020*″灌灌″ + 0.018*″重庆″ + 0.014*″南宁″ + 0.013*″美女″ + 0.011*″温州″ + 0.010*″妹子″ + 0.008*″盐城″ + 0.008*″农村″ bump posts, Chongqing, Nanning, beauty, Wenzhou, young girl, , countryside
Topic 28	0.008*″办理″ + 0.008*″申请″ + 0.006*″单词″ + 0.005*″小学″ + 0.005*″身份证″ + 0.005*″预约″ + 0.005*″孩子″ + 0.005*″登记″ conduct, application, word, primary school, personal ID, appointment, children, register
Topic 29	0.021*″垃圾″ + 0.018*″咖啡″ + 0.018*″宝″ + 0.010*″共享″ + 0.009*″上海″ + 0.009*″单车″ + 0.006*″星巴克″ + 0.006*″塑料″ garbage, coffee, treasure, share, Shanghai, bike, Starbucks, plastics
Topic 30	0.011*″地震″ + 0.011*″猪″ + 0.010*″王″ + 0.010*″内裤″ + 0.007*″按摩″ + 0.007*″颈椎″ + 0.006*″中国″ + 0.005*″近视″ earthquake, pig, king, underwear, massage, cervicle spine, China, myopia

Table III: Time-Trend of Topic Dispersion

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Age (month)	.0070***	.0060***	.0065***	.0061***	.0069***	.0074***	.0068***
	(.0001)	(.0001)	(.0003)	(.0004)	(.0001)	(.0002)	(.0001)
Initial Date	-.0005***	-.0003***	-.0005***	-.0005***	.0010***	-.0004***	-.0003***
	(.0001)	(.0001)	(.0001)	(.0001)	(.0001)	(.0001)	(.0001)
Repost Shock				.0118**			
				(.0048)			
Const	5.646***	5.824***	5.641***	4.988***	4.951***	6.146***	4.211***
	(.0439)	(.0441)	(.0455)	(.0375)	(.0393)	(.0474)	(.0239)
Influencer FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Main Topic FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	No	Yes	Yes	No	No	No
R-sq (overall)	.787	.734	.787	.787	.785	.798	.771
Obs	42056	42056	42056	42056	36434	42056	42056
Groups	1002	1002	1002	1002	885	1002	1002

* = 0.10, ** = 0.05, *** = 0.01

Table IV: Time-Trend of log Advertisements Share

	CNN	RNN	Keywords	CNN
	(1)	(2)	(3)	(4)
Age (month)	.0323***	.0328***	.0266***	.0329***
	(.0009)	(.0009)	(.0005)	(.0010)
Initial Date	.0014***	.0013***	-.0026***	.0010***
	(.0005)	(.0005)	(.0002)	.(0005)
Const	-5.468***	-5.331***	-3.283***	-5.153***
	(.2243)	(.2538)	(.1197)	(.2644)
Influencer FE	Yes	Yes	Yes	Yes
Main Topic FE	Yes	Yes	Yes	Yes
R-sq (overall)	.479	.476	.496	.483
Obs	41959	41959	41959	36006
Groups	1002	1002	1002	875

* = 0.10, ** = 0.05, *** = 0.01

Table V: Event Study of the Original Tag Shock

Dependent Variable:	$ln(Dispersion_{it})$				$ln(AdvShare_{it})$	
	(1)	(2)	(3)	(4)	(5)	(6)
Post Original Shock	.0358***	.0609***	.0393***	.0707***	.4316***	.5366***
	(.0106)	(.0145)	(.0118)	(.0162)	(.1126)	(.1561)
Age (month)	.0058***	.0001	.0046***	-.0029***	.0554***	.0474
	(.0014)	(.0039)	(.0016)	(.0044)	(.0163)	(.0447)
Initial Date	-.0001	-.0002	-.0001	-.0001	-.0021	-.0036
	(.0001)	(.0002)	(.0001)	(.0002)	(.0014)	(.0023)
Const	5.893***	5.852***	6.054***	6.036***	-3.477***	-3.747***
	(.0605)	(.0996)	(.0658)	(.1037)	(.5014)	(.7779)
Influencer FE	Yes	Yes	Yes	Yes	Yes	Yes
Main Topic FE	Yes	Yes	Yes	Yes	Yes	Yes
R-sq (overall)	.865	.885	.822	.848	.591	.656
Obs	8690	4592	8690	4592	8727	4622
Groups	948	948	948	948	948	948

* = 0.10, ** = 0.05, *** = 0.01

Table VI: Imbalance Check of Coarsened Exact Matching

	L1	mean	$\|\|t\|\|$	min	25%	50%	75%	max
Unmatched Data:								
Ave. Euclidean Dist.	.13049	-3.1982	1.02	9.3067	-11.467	-11.782	12.225	-408.42
Adv Share	.23882	.01524	8.90	0	-.01723	-.01333	.02153	.10127
Ave. Clicks	.09161	-3286.7	4.25	128.02	-2637.1	-3144.7	-4749	0
Ave. Likes	.03814	-46.706	0.91	.58824	-8.2866	-19.112	-24.644	-22433
Matched Data:								
Ave. Euclidean Dist.	.1171	-.7435	0.50	9.307	-5.093	-3.762	9.399	-62.98
Adv Share	.2425	-.0088	2.66	0	-.0171	-.0161	.0026	.0119
Ave. Clicks	.0330	110.2	0.10	128.0	188.2	9.209	40.03	0
Ave Likes	.0459	19.93	0.94	.5882	7.261	6.854	23.06	-1258

Table VII: Original Tag Shock: Coarsened Exact Matching and DID

Dependent Variable:	$ln(Dispersion_{it})$				$ln(AdvShare_{it})$	
	(1)	(2)	(3)	(4)	(5)	(6)
Post×Treat	.0522***	.0588***	.0556***	.0652***	.6645***	.6491***
	(.0059)	(.0081)	(.0067)	(.0090)	(.0392)	(.0516)
Post Shock	-.0028	-.0041	-.0015	-.0048	.0592***	.0414**
	(.0020)	(.0027)	(.0022)	(.0030)	(.0132)	(0.176)
Treatment Group	-.0062	-.0094	-.0086	-.0138*	-1.245***	-1.145***
	(.0054)	(.0075)	(.0060)	(.0084)	(.0356)	(.0483)
Age (month)	.0037***	.0035***	.0022***	.0020***	.0050***	.0043***
	(.0001)	(.0001)	(.0001)	(.0002)	(.0007)	(.0009)
Initial Date	-.0004***	-.0005***	-.0003**	-.0004**	-.0022***	-.0033***
	(.0001)	(.0002)	(.0001)	(.0001)	(.0008)	(.0010)
Const	5.894***	5.891***	6.138***	6.128***	-1.347***	-1.438***
	(.0610)	(.0840)	(.0803)	(.0934)	(.4007)	(.5414)
Influencer FE	Yes	Yes	Yes	Yes	Yes	Yes
Main Topic FE	Yes	Yes	Yes	Yes	Yes	Yes
R-sq (overall)	.809	.815	.737	.747	.542	.576
Obs	64792	33955	64792	33955	64880	34013
Groups	940	940	940	940	940	940

Table VIII: Compare Reactions to the Arrest of Meng Wanzhou

	DV: fraction of Meng-related Articles			
	(1)	(2)	(3)	(4)
Post*New	-.0030***	-.0032***	-.0026**	-.0036***
	(.0011)	(.0011)	(.0010)	(.0014)
Post	.0123***	.0123***	.0094***	.0143***
	(.0005)	(.0005)	(.0005)	(.0006)
New	.0000	-.0000	.0001*	.0007
	(.0001)	(.0001)	(.0000)	(.0005)
Const	.0002***	.0001***	.0001*	.0022***
	(.0000)	(.0000)	(.0000)	(.0002)
R-sq	.014	.014	.010	.013
Obs	50943	50058	47019	50943
# Accounts	910	895	840	910

* = 0.10, ** = 0.05, *** = 0.01

Table IX: Time-Trend of Likes-Clicks Ratio at Different Quantiles

	Whole Data				Beginning Two Years			
	Total	Low	Mid	High	Total	Low	Mid	High
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Age (month)	-.0024***	-.0051***	-.0015***	-7.07e-4**	-.0060***	-.0101***	-.0058***	-.0023***
	(.0003)	(.0003)	(.0003)	(.0003)	(.0008)	(.0010)	(.0007)	(.0008)
Initial Date	.0004**	.0006***	.0004***	.0002	.0011***	.0017***	.0012***	.0007***
	(.0002)	(.0002)	(.0001)	(.0002)	(.0002)	(.0002)	(.0001)	(.0002)
Const	-6.358***	-6.357***	-6.430***	-6.242***	-6.522***	-6.705***	-6.633***	-6.377***
	(.0871)	(.1210)	(.0712)	(.0929)	(.1007)	(.1323)	(.0867)	(.1141)
Influencer FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Main Topic FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-sq (overall)	.730	.639	.718	.731	.827	.717	.826	.818
Obs	17607	22018	22013	17607	9752	11994	11989	9752
Groups	562	562	563	562	562	562	563	562

* = 0.10, ** = 0.05, *** = 0.01

Appendix C. Numerical Simulation

A Numerical Example of Proposition 1

Assume that consumers match value to the influencer, $u_i \sim N(5, 1)$. I assume that $\Delta = 10$, $\theta_0 = 0.012$ and $\lambda_L = 0.67$.[27] Note that in this case, Assumption 1 does not hold if $\gamma > 0.87$.

Figure IX and X show the simulation result of optimal horizontal design of content x_t and amount of advertisements a_t respectively, for depreciation rate $\gamma = 0.9, 0.7$ and 0.5. First, for all values of γ, both the horizontal design x_t and amount of advertisements a_t unambiguously monotonically increase over time, which echo the results in Proposition 1. This is true also for $\gamma = 0.9$ case, which violates Assumption 1.

Comparing the growth curves of different γ in Figure IX, it can be seen that a greater γ (fewer subscribers carry out from past) associates to an earlier transition to broad content. This is consistent to the prediction in the Lemma 4, such that the lower γ is, the longer the influencer stays in the most niche stage. Also, the transition process takes time due to the accumulation of past subscribers and there exists a "transitional stage" between the most niche ($x_t = 0$) and most broad ($x_t = 0.5$) design, which is due to the accumulation of relatively low match valued type A audiences from the past. Figure X shows that a higher γ corresponds to a less advertisements, especially after the influncer start to transit to broad content design. It is interesting to note that during the transition process (for example, $t = 8$ to $t = 10$ in the $\gamma = 0.7$ curve of Figure X), advertisement grows slower than the "most niche" periods. The intuition is that in this time, it is most efficient for the influencer to set a most broad content to attract the new audiences. Therefore the influencer has higher incentive to adjust the component of his subscriber base so to make the two types more balanced. This is why during this time, the influencer restrains the advertising level so to replenish more type B audiences to his subscriber base.

In the main model, we assume that the influencer is myopic for tractability consideration. Here I also provide a numerical simulation when the influencer is forward looking. I assume that at each period of time, the influencer maximizes the summation of all the future payoffs.[28] All the

[27] $\theta_0 = 0.012$ and $\lambda_L = 0.67$ implies that at $t = 0$, the probability that an influencer produces a bad content, $\xi_0 = 0.325$.
[28] This is convergent when $\gamma > 0$.

parameters are the same as above and $\gamma = 0.7$. Figure XI and XII display the result. It is intuitive that the level of advertisements in the forward looking scenario is lower than the myopic case, as the marginal benefit of acquiring one new subscriber is greater for a forward looking influencer. This causes the reversal in the dynamic of one period profit as shown in Figure XII: in the very beginning the one period profit of myopic influencer is greater as as the forward looking influencer ”over-invest” in the subscriber base. But soon the static profit of forward looking influencer exceeds the myopic one.

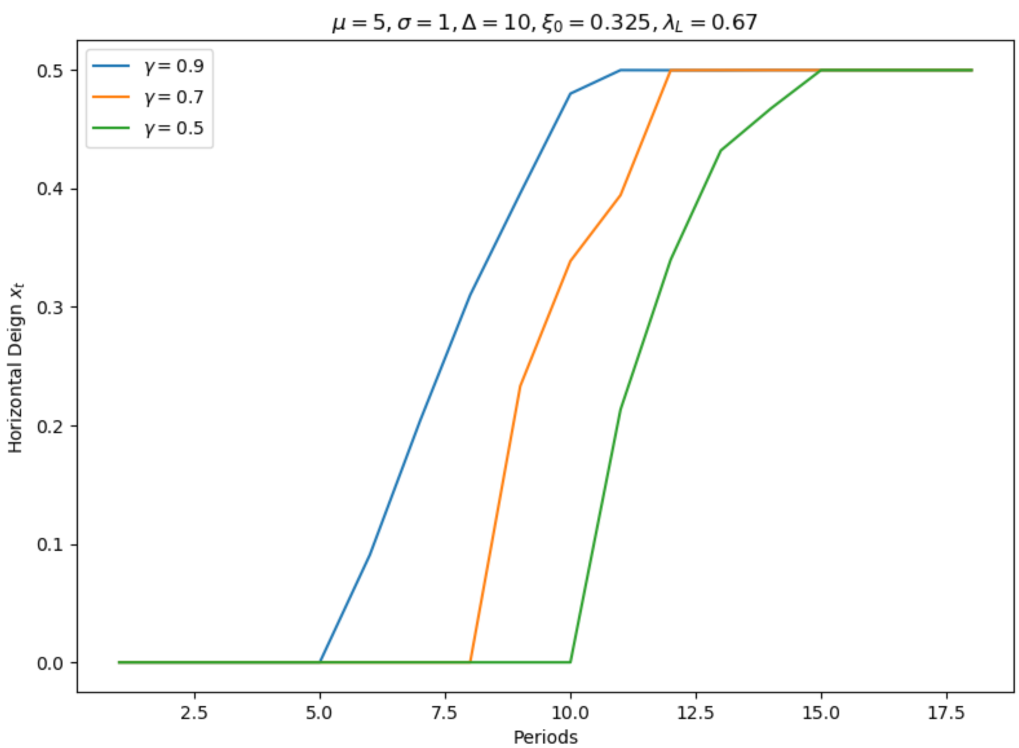

Figure IX: Numerical Simulation: Dynamics of Horizontal Design

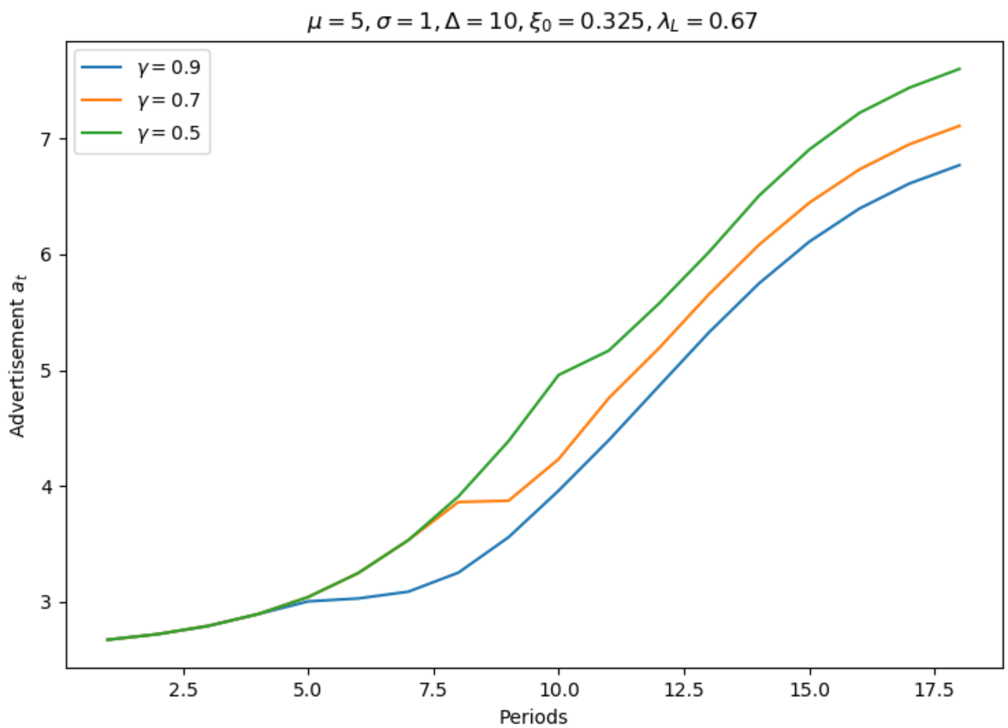

Figure X: Numerical Simulation: Dynamics of Advertisement

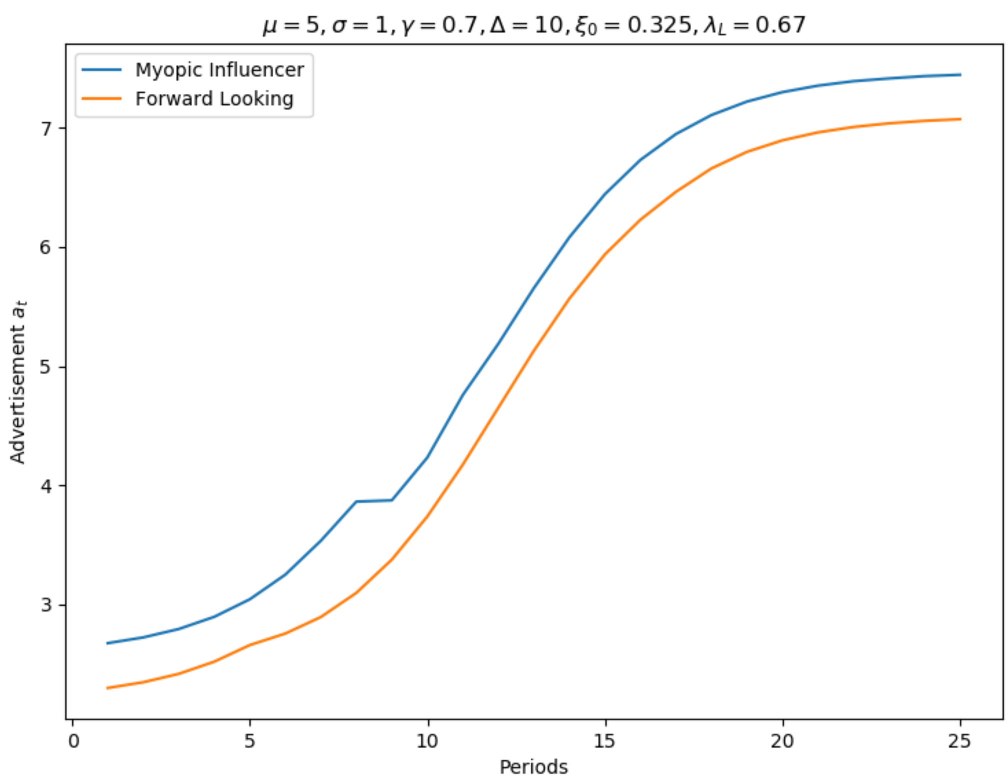

Figure XI: Comparison of Forward Looking and Myopic Influencers: Advertisement

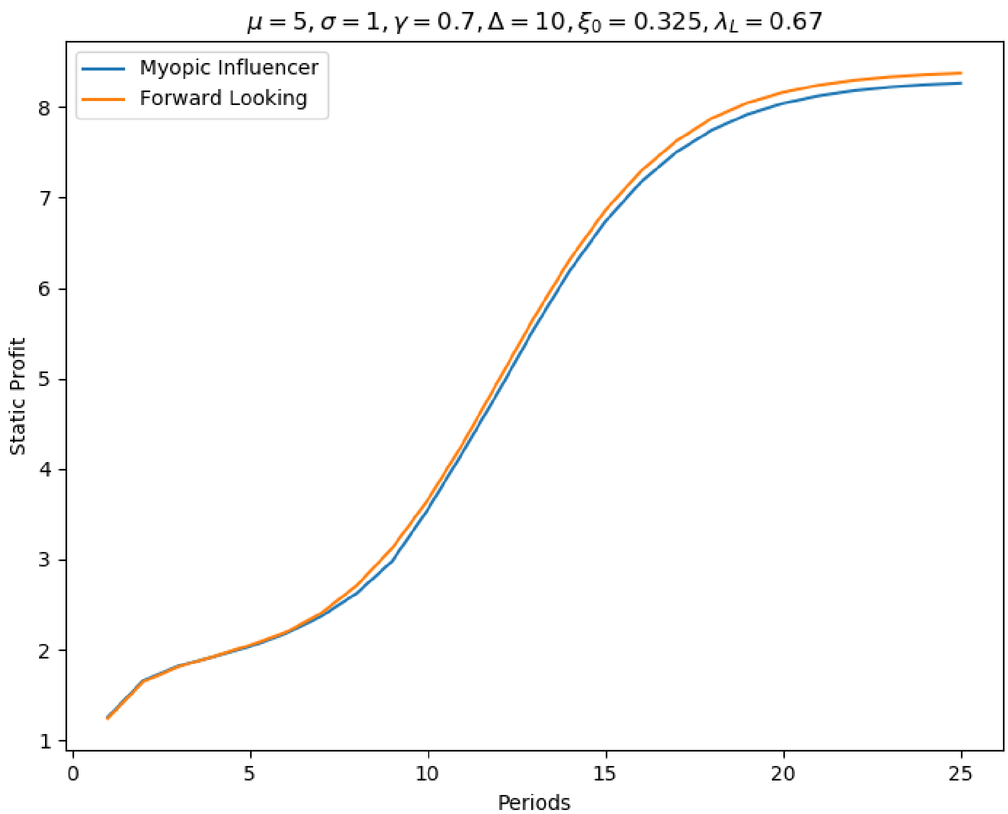

Figure XII: Comparison of Forward Looking and Myopic Influencers: Static Profit

Propensity Score Matching in the Original Tag Shock

I present the result using propensity score matching method (Rosenbaum and Rubin, 1983) as a robustness check, which suggests the use of the probability of receiving treatment (original tag function) to find a best control, j, for each influencer i at there respective period of treatment $\hat{t}_i$. We first identify the probability (the propensity score) of receiving the original tag invitation with a probit model:

$$Pr(OT_{it} = 1) = F(X_{it-1}) \tag{12}$$

where X is a vector of observable covariates, including average clicks, likes, number of articles, average content dispersion and advertising share. Next, for each influencer i who receive the treatment at $\hat{t}_i$ with propensity score $p_{i,\hat{t}_i}$, we find the influencer j at time t_j, such that the propensity score p_{j,t_j} is closest to $p_{i,\hat{t}_i}$ such that influencer j exists and does not receive the original tag shock within one year of time around t_j ($[t_j - 6, t_j + 5]$). We use the data of influencer j at time $[t_j - 6, t_j + 5]$ as the control group for influencer i.

With the propensity score matched control group for each influencer's treatment, we now can apply a standard difference-in-difference approach to estimate the impact of original tag function on the content design. The econometric specification is as follows:

$$y_{it} = \alpha_i + \alpha_k + \beta_1 Post_{it} + \beta_2 Treat_i + \beta_3 Post_{it} \times Treat_i + \beta_4 Age_{it} + \beta_5 InitialDate_i + \epsilon_{it}$$

for influencer i, time t and main topic k. y_{it} is either the topic diversity, $ln(Dispersion_{it})$ or portion of advertising posts, $ln(AdvShare_{it})$. $Post_{it}$ is a dummy variable that equals to 1, if influencer i is in the treatment group and the time is after the first period he publishes an original tag content $\hat{t}_i$, or if (j, t_j) is the control group of $(i, \hat{t}_i)$ and $t \geq t_j$. $Treat_i$ is a dummy variable that equals to 1 for treatment group observations. Age_{it} is the age (in month) of influencer i at time t, and we control for the influencer i FE by α_i and main topic k FE for publications in $\mathcal{A}_{it}$.

Table X shows the result. Same to the event study, columns (1) and (2) uses average Euclidean distance as the measure of topic diversity, while (3) and (4) uses the standard deviation. Columns

(1), (3) and (5) use 6 months of pre-trend and post-trend, abd (2), (4) and (6) use 3 months of data. I find that compared to the control group, the original tag shock increases 3.7% to 5.5% of content dispersion, which equivalents to the increment of topic dispersion of 4.9 - 7.8 months. On the share of advertisements, receiving the original tag function increases the amount of advertisement by 37% to 43%, which equivalents to about 15-18 months of the increment in advertising posts.

Table X: Original Tag Shock: Propensity Score Matching and DID

Dependent Variable:	$ln(Dispersion_{it})$				$ln(AdvShare_{it})$	
	(1)	(2)	(3)	(4)	(5)	(6)
Post×Treat	.0365***	.0477***	.0411***	.0545***	.4273***	.3720***
	(.0132)	(.0129)	(.0138)	(.0140)	(.0806)	(.0850)
Post Shock	-.0113	-.0060	-.0133	-.0089	-.0088	.0343
	(.0094)	(.0068)	(.0099)	(.0073)	(.0357)	(0.411)
Treatment	-.0625***	-.0647***	-.0693***	-.0739***	-.4626***	-.3773***
	(.0148)	(.0165)	(.0153)	(.0176)	(.0940)	(.1099)
Age (month)	.0079***	.0081***	.0069***	.0070***	.0230***	.0235***
	(.0015)	(.0016)	(.0016)	(.0018)	(.0059)	(.0061)
Initial Date	.0002***	.0002***	.0001***	.0001***	.0002*	.0001
	(.0000)	(.0000)	(.0000)	(.0000)	(.0001)	(.0001)
Const	4.941***	4.919***	5.148***	5.146***	-4.930***	-5.184***
	(.0420)	(.0512)	(.0486)	(.0548)	(.2726)	(.2815)
Influencer FE	Yes	Yes	Yes	Yes	Yes	Yes
Main Topic FE	Yes	Yes	Yes	Yes	Yes	Yes
R-sq (overall)	.228	.244	.212	.233	.183	.191
Obs	20028	10262	20028	10262	20065	10292
Groups	984	984	984	984	984	984