

Effect of Popularity Shocks on User Behavior

Omkar Gurjar¹, Tanmay Bansal², Hitkul Jangra², Hemank Lamba³, Ponnurangam Kumaraguru¹

¹International Institute of Information Technology, Hyderabad, India

²Indraprastha Institute of Information Technology, Delhi, India

³Carnegie Mellon University, USA

omkar.gurjar@students.iiit.ac.in, {tanmay17268, hitkuli}@iiitd.ac.in, hlamba@cs.cmu.edu, pk.guru@iiit.ac.in

Abstract

Users often post on content-sharing platforms in the hope of attracting high engagement from viewers. Some posts receive unusual attention and go “viral”, eliciting a significant response (likes, views, shares) to the creator in the form of *popularity shocks*. Past theories have suggested a sense of reputation as one of the key drivers of online activity and the tendency of users to repeat fruitful behaviors. Based on these, we theorize popularity shocks to be linked with changes in the behavior of users. In this paper, we propose a framework to study the changes in user activity in terms of frequency of posting and content posted around popularity shocks. Further, given the sudden nature of their occurrence, we look into the survival durations of effects associated with these shocks. We observe that popularity shocks lead to an increase in the posting frequency of users, and users alter their content to match with the one which resulted in the shock. Also, it is found that shocks are tough to maintain, with effects fading within a few days for most users. High response from viewers and diversification of content posted is found to be linked with longer survival durations of the shock effects. We believe our work fills the gap related to observing users’ online behavior exposed to sudden popularity and has widespread implications for platforms, users, and brands involved in marketing on such platforms.

Introduction

Recently, social media platforms have emerged or transformed themselves to focus more on content creation and sharing, e.g., TikTok, Instagram, Twitch, YouTube, etc. These social media platforms, focusing on content/multimedia sharing, have enabled users to express themselves in unique ways (text, photos, videos, etc.) to their followers (subscribers). To continue content creation and also engagement, most of the platforms have also launched creators’ funds and also allow content creators (users) to get incentives/money to create such content (Kopf 2020; Veissi 2017). With social media content creation becoming an alternate source for revenue generation, users are also focusing on creating exciting content and eliciting attention and thus engagement from other users.

Users who become popular on these social media platforms are often termed as “influencers” or “micro-

celebrities” (Veissi 2017; Freberg et al. 2011; Jin, Muqadam, and Ryu 2019). Influencers, due to their popularity, have a broad reach and have been studied in the past on swaying/forming attitudes about consumer purchase intention (Lim et al. 2017), brand’s image (Hermanda, Sumarwan, and Tinaprillia 2019) and perceived uniqueness (De Veirman, Cauberghe, and Hudders 2017), along with even dietary behavior in children (Smit et al. 2020). Influencers are also often contacted by different brands for endorsing their products (Veissi 2017; Uzunoğlu and Kip 2014). Many studies have been done in understanding why certain content and users who post them become popular (Figueiredo et al. 2014; Gabielkov et al. 2016; Figueiredo, Benevenuto, and Almeida 2011; Mazloom et al. 2016). However, there has been little or no study on addressing how viral users respond to their newly achieved popularity.

On content sharing platforms, receiving sudden popularity due to specific content (or a series of content getting viral) can be termed as *popularity shocks*. Popularity shocks can be characterized as a sudden increase in feedback (i.e., views, likes, etc.). Previously, popularity shocks have been studied towards Wikipedia pages because of an associated event (Zhang et al. 2019; Keegan and Brubaker 2015), and Github repositories due to being highlighted by the platform (Maldeniya et al. 2020). However, the effect of popularity shocks on users’ content creating behavior has not been studied in detail. Similarly, much work has been done in predicting posts that will go viral or will become popular using initial dynamics (Figueiredo 2013; Zhao et al. 2015; Weng, Menczer, and Ahn 2013). However, little work has been done in analyzing the after-effects of a post becoming viral or a user becoming popular. *Do users become more active on the platform after getting popular? Do users alter their content or stick to the content that made them popular? How long does the popularity shock last? Is popularity short-lived, or can it be long-term based on how user conducts themselves?* Answering these questions could have wide-reaching implications for all three - the users, potential brands seeking influencers to partner with, and also the social media platform itself. Studying users’ response to popularity shock can be insightful for (a) users, who want to continue engagement, (b) brands, for identifying new influencers which align with their values, and (c) social media platforms, for guiding new popular users on specific inter-

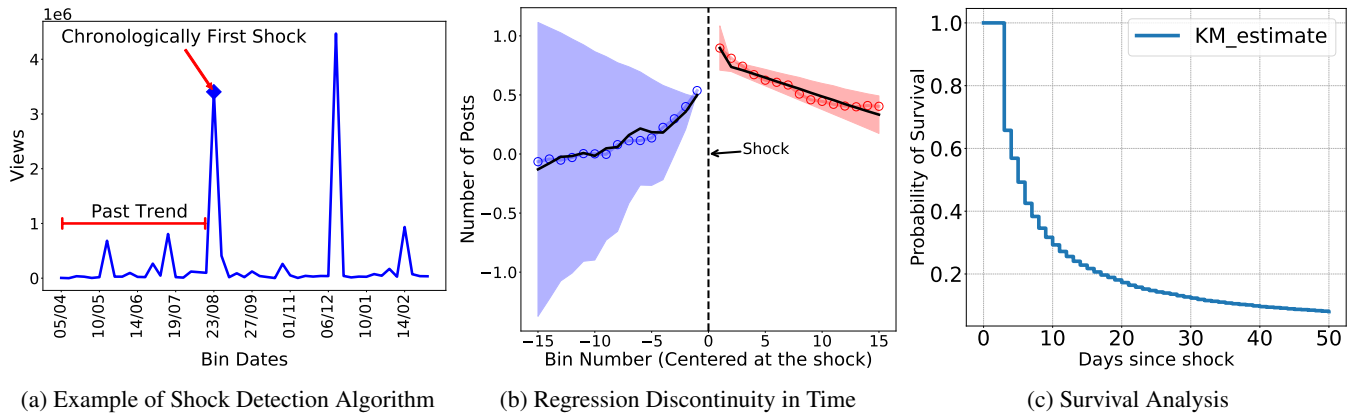


Figure 1: Our work discusses (a) detecting points of sudden increase in response known as *popularity shocks* on users’ timelines; (b) Quantifying behaviour change due to popularity shock in terms of change in posting frequency using RDiT(Regression Discontinuity in Time); (c) Short-lived survival duration of effect of shocks and factors affecting it.

ventions that can be related to education, design changes or guidelines.

We ground our work in sociological theories related to social reinforcement and a sense of reputation. A reputable theory in the field of behavioral psychology has been *Operant Conditioning* (Skinner 1938). Under this theory, an activity that earns rewards prompts an individual to repeat that activity, and similarly, an activity that earns punishment makes the individual more inclined to repeat that activity. In our context, if we treat receiving popularity, which is quantified with high engagement from the community on users’ content, as positive feedback (or reward), the user ideally will keep repeating the same behavior. Alternatively, if the user received a popularity shock in a negative context, i.e., they were a recipient of a firestorm (Lamba, Malik, and Pfeffer 2015), they might stop posting similar content. We also draw on the theoretical work carried out in a more specific context of online communities (Kollock et al. 1999; Rheingold 1993). In one of the earliest analyses of an online community, Rheingold hypothesized that desire for prestige is one of the key motivators for individuals’ contribution to the community. Kollack re-emphasized this (Kollock et al. 1999), highlighting that increased reputation is one of the three reasons for individuals to contribute content on online platforms. Contextualizing this in our work, popularity shock can be viewed as a signal of increasing reputation and might prompt users to continue contributing to the platform. Though these theories were proposed some time ago, rigorous empirical evaluation/validation of these theories in the context of popularity on online social media platforms have not yet been conducted.

In this paper, we study how do users’ behavior changes after a popularity shock in terms of (a) frequency of posting, (b) the content, itself and (c) how long do they continue with their altered behavior. We first characterize what should be considered as a popularity shock and develop a method to identify popularity shocks from a user timeline. Using popularity shock as an intervention, we use causal inference techniques to examine the change in behavior from

pre-and-post popularity shock. Next, we study the change in the content posted by users under the effect of popularity shock. We leverage document embeddings (Le and Mikolov 2014) to model the posted content mathematically. Finally, we investigate the expected duration for shock’s effect and its dependence on other factors using survival analysis techniques.

Data and Code: We released the anonymized version of our data available at: <https://precog.iit.ac.in/research/effect-popularity/>

Related Work

Since our work is related to users’ response towards increased attention, our related work flows from three main directions - (a) Effect of social feedback, (b) Attention Shocks and (c) Popularity/ Virality Prediction.

Effect of social feedback: Positive reinforcement or feedback has been a popular area of study among social scientists (Rushton and Teachman 1978; Pavett 1983; Saari and Näsänen 1989; Balcazar, Hopkins, and Suarez 1985; McMillen and Austin 1971). (Rushton and Teachman 1978) demonstrated through experiments on around 60 children through a bowling game that positive reinforcement led to improvement in altruistic behavior in children, while punishment led to the opposite. This framework has been studied extensively in various settings, such as effect of positive feedback on promoting safe behaviours in housekeeping (Saari and Näsänen 1989) and effect on compliance following transgression (McMillen and Austin 1971) as well as simulating motivations and future play of a brain training game (Burgers et al. 2015). In the domain of online world, however opposite effect has been observed in the case of low quality comments (Cheng, Danescu-Niculescu-Mizil, and Leskovec 2014), where it was observed that negative feedback prompted users to continue with writing low quality comments on news articles. Further, (Chen, Dhana-sobhon, and Smith 2008) how the community perception of helpfulness of online reviews, influences consumer purchase decisions, and how this helpfulness vote is itself de-

terminated by evaluations of the same product by the community (Danescu-Niculescu-Mizil et al. 2009; Sipos, Ghosh, and Joachims 2014). Similar study has also been conducted for the effect of social feedback on weight loss community (Cunha et al. 2016).

Though there has been a lot of studies discussing social feedback, however very few have tried to characterize how do users or actors in turn respond to extremely high and sudden feedback in data-oriented fashion on a large-scale data.

Attention Shocks: Attention shocks are characterized as sudden attention being drawn towards a specific entity (any author/artefact on social media platform). Examples include, death of a celebrity leading to increased attention towards the celebrity’s wikipedia page (Zhang et al. 2019). (Maldeniya et al. 2020) use the lens of organisation change to study the dynamics of change in behaviour of contributors of a GitHub repository experiencing increased attention as a result of being listed on the trending page. On Wikipedia, (Zhang et al. 2019) observe increased participation of new comers and study collaborator dynamics on pages in times of shock detected through Google Trends, while (Zhang et al. 2017) look into the changes in collaborative behaviour of editors due to shock resulting from imposition of censorship in mainland China. Other works like (Keegan, Gergle, and Contractor 2013) study similar changes in case of breaking news articles on Wikipedia. (Lamba, Malik, and Pfeffer 2015) analyse shocks in form of sudden bursts of negative attention towards controversial events called ‘firestorms’, and use Twitter data to characterize the size and longevity of these firestorms.

Other works study the effect on online network structures under shocks. (Keegan and Brubaker 2015) suggest the formation of complex but temporary collaboration networks of users during increased editing activity on Wikipedia page of a diseased person and study their dynamics. Further, (Keegan, Gergle, and Contractor 2012) introduce a method of capturing collaboration structure of co-authors of a Wikipedia articles and highlight the difference between such networks for breaking news articles, as compared to traditional ones based on pre-existing knowledge.

Though attention shocks have been studied on online social media platform, to the best of our knowledge, our work is the first attempt to study the behaviour of users whose posts goes viral (i.e. the user who gets the shock). A minor characteristic that differs us from other studies is that we are looking at shock as a sudden virality of the post, and the virality of the post is mostly algorithm-driven (i.e. probably a mixture of recommendation algorithm and “rich-gets-richer” theory). In comparison, other studies looked at shock which was more exogenous i.e. appearance on GitHub trending page or death of a celebrity. Lastly, there are inherent differences in nature of platforms being studied. While Github and Wikipedia are collaborative platforms where users are often driven by non-monetary motivations such as reputation and collective identity (Maldeniya et al. 2020), users on such content sharing platforms are driven by monetary causes and for self-satisfaction. Thus there is clear distinction in intent of use, due to which we can expect difference in user behaviour as well.

Though it is not highly aligned with our work, however there has been significant amount of work done for predicting if a post is going to get popular or not, and hence we mention about some of the efforts done to solve that problem.

Content Virality Most work in this domain is focused on predicting and characterising virality of online content. (Figueiredo 2013) understands popularity trends for online user generated content (UGC) in the form of online videos, and proposes a prediction model based on extremely random ensemble tree to predict the popularity trends for Youtube videos. The SEISMIC model proposed by (Zhao et al. 2015) predicts the final number of reshares a post will receive based on the past history. The problem is modelled as predicting the final size of an information cascade and performance is validated on a month of Twitter data. Other models like (Wang, Bansal, and Frahm 2018), (McParlane, Moshfeghi, and Jose 2014) have tackled the problem of virality prediction on Twitter and Flickr respectively.

Other works are inclined more towards characterizing virality and viral content. (Weng, Menczer, and Ahn 2013) studies the virality and diffusion of memes on online networks. (Mazloom et al. 2016) seeks to identify features in posts which are related to its popularity using a multi-modal approach. (Figueiredo, Benevenuto, and Almeida 2011) aims to characterize and understand popularity growth of videos, and what kinds of mechanisms contribute towards popularity. The work also mentions presence of sudden bursts of popularity on top listed videos.

Theory and Research Questions

Kollock (Kollock et al. 1999) hypothesized that there are three significant reasons for users to keep on contributing to the social community - (a) anticipated reciprocity; user is generally motivated to contribute or stay as an active participant in online communities in the expectation that the user will receive helpful information when they are in need, (b) sense of efficacy; the users might contribute information because they are rewarded with the sense that they contributed something to the community (Bandura 1977). The efficacy can also result in the self-belief that they have a high impact on the community, hence providing the validation of their self-image as an efficacious person, and (c) Reputation; most users want recognition for their contributions or their efforts. As quantified by the number of unique impressions of their content, popularity validates their content. This can be seen as an increase in reputation for the user based on the high number of people that follow or subscribe to them. On the lines of Kollock, we hypothesize that receiving a popularity shock (i.e., increase in reputation) will prompt users to increase their activity on online social media platforms.¹ Therefore, we ask the following question:

RQ1. [Engagement Response to Popularity] *Do users increase their posting behavior after receiving popularity shock?*

¹In this work, we discovered that the popularity shocks were positive, analysis can be done if this popularity instead was negative too.

Category	Hashtags	Unique Users
Animals	cats, dogs, pets	1666
Beauty/Makeup	beauty, makeup, naturalbeauty, skincare	3052
Craft/DIY	5_min_craft, craftchallenge, diycraft, easycraft	1128
Dance	dance, dancechallenge, dancekpop	2144
Education	careergoals, education, learning, mindpower	2429
Entertainment	entertainment	449
Fitness	fitness, fitnessgoals, gym, weightloss, workout	3911
Food	food, foodislove, foodrecipe, healthyfood, myrecipe	2815
Funny	comedy, funny, meme	2632
Health	wellness	558
Motivational	advice, inspirational, lifehacks	2146
Music	hiphop, music	1323
Pranks	prank	667
Sports	cricket, football, sports, tennis	2341

Table 1: Number of unique users for each category (arranged in alphabetical order of Category).

Another social theory framework that fits very well with our setting is that of *operant conditioning* (Skinner 1938). Skinner theorized that the reward for action leads the agent to keep on performing the same action in anticipation of reward, and a punishment hinders the user’s propensity to take that action. Again, operationalizing reward as the popularity shock, we can hypothesize that users who received popularity shock will continue with the same behavior that earned them the reward even in our setting. This brings us to the following research question:

RQ2. [Content Response to Popularity] *Do users alter their content post receiving popularity shock?*

In network science, the transition of network states and dynamics due to an external event has been a topic of interest (Zhang et al. 2019; Maldeniya et al. 2020; Keegan and Brubaker 2015). (Malik and Pfeffer 2016) argue that some of the network transitions, and along with it changes in user behavior in these networks, are more permanent. Moreover, some studies argue that networks bounce back after the event, and normal communication ensues (Lamba, Malik, and Pfeffer 2015). In our setting, we were interested in understanding how long the popularity shock lasts.

RQ3. [Longevity of Effect] *How long do the effects of popularity shock last?*

For users who receive the popularity shock, it is imperative to understand what users can do or how they should maintain their activity that can prolong the shock’s effect. Therefore we ask the following question:

RQ4. [Sustained Shock Effect] *What type of activity characterizes long-term sustainability of effects of popularity shock?*

Data Collection

Background: We collect data from popular multimedia sharing social media platform.² On the platform, users can post multimedia content (images/videos) along with an associated caption. Depending upon the privacy setting of the

²name of the platform suppressed to retain anonymity and non-public API access.

post and the user’s profile, other users can view their content and engage with the content using platform-provided mechanisms such as liking the content, commenting on the post, or resharing the post. By liking, a user can express their positive response or acknowledgment, sharing works to amplify the reach of content, and viewers can also express their opinions in the form of comments. Like all other social networking platforms, the social platform under study also provides functionality that allows users to ‘follow’ other users on the platform. Besides this, the platform can also grant a special ‘verified’ status to specific users based on their strong influence on the platform or in the real world. Though we study a specific platform, we believe that a similar methodology can be applied to any social media platform with similar mechanisms in place. A cross-platform study on measuring this behavior and ensuring generalizability is one of the promising future directions of this work.

Data Collection: We identify 14 generic categories related to commonly posted content on the platform. From the list of these 14 categories, we curated a list of 43 popular hashtags. The hashtag selection was made keeping the goal of generalization in mind, and hence no hashtags related to specific entities (e.g., #ronaldocr7) were considered. The selected categories and hashtags are described in Table 1. Approximately 4,000 posts per hashtag were collected, coming from 21,224 unique users. Next, we collect posts liked by these users and add the authors of the posts to our dataset to minimize any sampling bias due to the collection strategy (which might be due to bias in the platform’s search functionality)³. Finally, we had a total of 33,490 users. We collected the entire timeline of these users and filtered out users who had less than 200 posts in their entire lifetime to ensure we had substantial data for our analysis.

Following the filtering, our final dataset contains a total of 30,969 users. We describe the data statistics in Table 2 along with distribution of number of posts across users in Figure 2.

³Data collection was done when the first and second authors were students at their respective institutes

For each post, we collected the following details of the post - (a) post id, unique identifier for the post, (b) timestamp of when the post was published, (c) caption of the post, (d) number of views the post received, (e) number of likes the post received, (f) number of times the post was reshared, (g) number of comments the post received and (h) user information - all key statistics such as name, bio, etc. of the user who created the post.

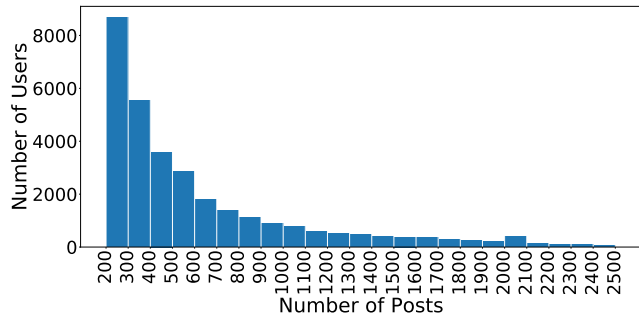


Figure 2: Distribution of Users’ Total Posts (Follows Power Law).

Number of Users	30,969
Number of Posts	18,911,417
Timestamp of First Post	7 th Jan 2015
Timestamp of Last Post	31 st Dec 2021

Table 2: Dataset Details.

Detecting Popularity Shocks

To answer any of the research questions mentioned above, we first need an algorithm that can identify popularity shocks from a user’s timeline. Before going into the details of the algorithm, we describe the assumptions we made to define popularity shock.

- We use the number of views as a proxy for popularity. Views give a more objective metric of the reach or engagement as it is implicit, unlike other metrics such as the number of likes, shares, or comments which require explicit action from the audience.
- A user might receive multiple popularity shocks throughout their career. However, we only study effects due to the chronologically first shock the users receive. We do not consider later shocks as the user would have already experienced some popularity until that point. In this paper, we want to characterize the effect of the first popularity shock when the sudden growth in popularity is unexpected for the user.

A desired shock detection algorithm should detect a sudden percentage increase in views of the user, we should also account for absolute thresholds to avoid false positives caused by the base effect. The first natural candidates for the task are time-series anomaly detection algorithms like Z-score (Catalbas et al. 2017) or Facebook’s Prophet (Taylor and Letham 2018). However, these algorithms consider

time-series signals in isolation and do not account for global thresholds. To curb this, we also experiment with a custom algorithm as presented in Algorithm 1.

We preprocess the timeline by binning the posts, where each bin is a period of consecutive D (bin size) days. For each user, we iterate over the bins in chronological order (Line 6). We maintain a running average of views of all the bins encountered so far (Line 13). Once we have processed the bin (i.e., no more posts need to be counted for that bin), we compute the ratio of views of the bin to the running average of bins before it. Note that we ignore bins with no posts while computing the running average. This ratio needs to be higher than a ratio threshold θ for it to be considered a shock candidate. To account for the cases where the running average is very low, we also consider the difference between current views and the running average, which needs to be greater than the base threshold η . Therefore, the first bin satisfying these two conditions is classified as the popularity shock for the user. If no point satisfies these conditions, we consider the user is without a popularity shock.

Ideally, keeping consistent with our shock assumptions, we want to capture the first post at which user perceives they might have gotten popular. To evaluate our detection algorithm, we conduct a verification experiment. We solicited annotations from long-term social media users, who were asked to independently look at the view timeline of 100 users and mark what they deem as the first instance a user would have felt popularity shock. The annotators had a Fleiss’ Kappa score (Fleiss 1971) of 0.60, which indicates moderate agreement (Landis and Koch 1977). Each sample was annotated by 3 annotators, and a clear majority was received in 93 instances out of 100. We compared the efficacy of our proposed approach with baselines of z-score and Prophet algorithm using the ground truth set. Predictions were obtained across a range of hyper-parameters for all algorithms, best achieved results are shown in Table 3. Our proposed shock detection algorithm performs the best and is used to detect popularity shocks in our further experiments.

Algorithm	Accuracy
Z-Score	23.6%
Prophet	42.5%
Proposed	66.6%

Table 3: Shock detection accuracy against the manually annotated ground truth. Proposed algorithm outperforms other baselines.

The percentage of users we can discover having a popularity shock with different values of θ and η is presented in Figure 3. Note that, we report results with hyper-parameters $D = 1, \theta = 50, \eta = 1.5M$, unless specified. However, we experimented with multiple values of θ and η , and results stay consistent over reasonable values of these thresholds.

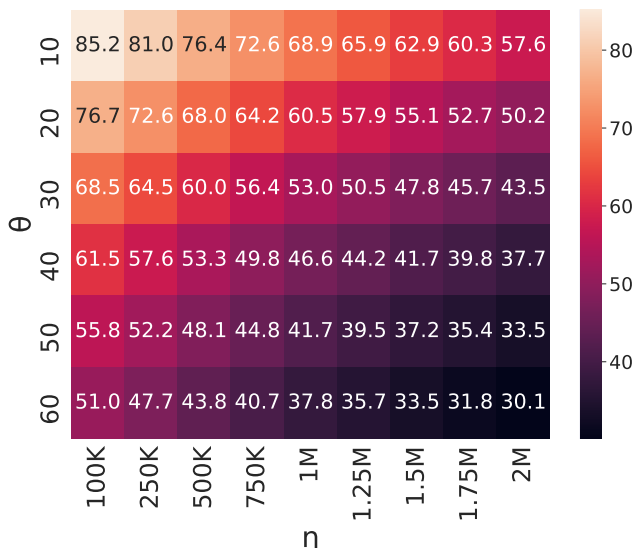


Figure 3: Heatmap representing percentage of users detected with a shock for different values of θ and η for $D = 1$. θ is the minimum ratio of views in the bin to the running average, while η is the minimum difference between the two, for detecting shocks.

Effect of Popularity

RQ1 seeks to quantify the change in posting frequency of a user due to the shock received. We do this using a causal inference technique called Regression Discontinuity in Time (RDiT) (Hausman and Rapson 2018).

Regression Discontinuity Design (RDD): Introduced by (Thistlethwaite and Campbell 1960), RDD is a quasi-experimental technique to measure the effects of a treatment or intervention. The population receives the treatment having the value of running variable X above a certain threshold known as the ‘cut-off’ point, and data is checked for any jumps or discontinuities in the outcome variable Y around the cut-off. Previously, RDD has been widely used in fields

Algorithm 1 Shock Detection Algorithm

```

1: function DETECTSHOCK(posts, D,  $\theta$ ,  $\eta$ )
2:   bins  $\leftarrow$  bin_data(posts, bin_size=D)
3:   shock  $\leftarrow$  -1
4:   n  $\leftarrow$  length(bins)  $\triangleright$  Total number of bins
5:   run_avg = views(bins[1])
6:   for i in 2 to n do
7:     ratio  $\leftarrow$  views(bins[i])/run_avg
8:     diff  $\leftarrow$  views(bins[i]) - run_avg
9:     if ratio  $\geq$   $\theta$  and diff  $\geq$   $\eta$  then
10:      shock  $\leftarrow$  i
11:      break  $\triangleright$  break at the earliest shock
12:     end if
13:   run_avg  $\leftarrow$  mean(views(bins[1:i]))
14: end for
15: return shock  $\triangleright$  if shock is -1, no shock found

```

such as Economics (Lee and Lemieux 2010) and Psychology (Cook 2008). Specifically, on social media studies, RDD has been analyzed previously to quantify the effect of obtaining a GitHub badge on users’ posting frequency (Okta, Taylor, and Jensen 2010), on the effect of the introduction of Facebook ‘‘People you may know’’ feature (Malik and Pfeffer 2016), and also on the effect of averaging rounding stars on Yelp (Li 2013).

Acknowledging that time being the running variable might cause some of the assumptions of traditional RDD not to hold, we use a variation of the RDD framework called **Regression Discontinuity in Time (RDiT)** proposed in (Hausman and Rapson 2018), in which time is the running variable and a fixed point in time is taken as the threshold. RDiT conceptually differs from the regular RDD on the following fronts:

- While RDiT aligns with the ‘discontinuity at cut-off’ interpretation of RDD, the ‘local randomization’ interpretation may not hold as the time assignment can not be taken as entirely random around the cut-off.
- Unlike RDD, sample size can not be grown arbitrarily with smaller bandwidths. Due to this, data points far from the cut-off need to be included, which can introduce biases due to changes in unobserved confounders over time.
- Including covariates becomes far more critical to control biases since the assignment of treatment and control groups is not entirely random around the cut-off.

Our methodology. To model our problem using RDiT, we define our running or forcing variable X as the bin index (signifying time) and outcome variable Y as the number of posts done by the user in the bin X . The shock bin is assigned index 0; subsequently, index $+i$ denotes the i^{th} bin after the shock, while the index $-i$ denotes i^{th} bin preceding the shock. The cut-off point c is $X = 0$, where the shock occurs. Then, treatment group is defined as $\{(X_i, Y_i) \text{ s.t. } X_i > 0\}$ and control group as $\{(X_i, Y_i) \text{ s.t. } X_i < 0\}$. We also control for the following covariates in our regression design:

- **Intensity of shock:** To account for variation in treatment, we control for the intensity of shock obtained in the preceding bin. The intensity is the value of the *ratio* variable for the bin as in Algorithm 1. We take the logarithm of this variable.
- **Age of User:** As receiving a popularity shock at different stages of users’ online life might have different effects. We control for the number of days since the user’s first post.

We then fit models separately on the two groups using regression. We only use W bins before and after the shock bin to fit the lines to avoid any effects of future shocks. On obtaining the equations of the two lines, their values at the cut-off point are predicted, which are used to calculate the discontinuity at the cut-off. Formally, let $Y_{t,0}$ and $Y_{c,0}$ be the values at the cut-off for the treatment and control lines respectively, then discontinuity at the shock d is given by $d = Y_{t,0} - Y_{c,0}$. From the equation, it can be seen that

a positive d corresponds to an increase in the frequency of posting after the shock as compared to before and vice versa.

Effect on Posting Frequency

We tried to estimate the effect of popularity shock on the posting frequency of user post-shock using RDiT. We quantified the intervention to occur at the time-point where we detected the popularity shock. Further, we count the total number of views that the user received each day before and after the shock. Note that this corresponds to setting $D = 1$ in Algorithm 1.⁴ In Figure 1b, we visualize the effect on posting frequency. The x-axis clearly shows the time before and after the shock. To aggregate the effect across all users, we compute the number of posts done by the user each day subtracted by the average number of posts done by the user in the past 15 days (this is done to maintain a consistent scale across users). Then, the average is taken across all users (including covariates) and curves as fit. The vertical dashed line shows the day on which popularity shock was observed. As mentioned above, we fit two linear regression models.⁵ The first model is for the average number of posts done before the popularity shock, and the second one is for the average number of posts after the popularity shock. We see a significant difference between the intercept and the slope for both the regression models. The discontinuity at shock (d) estimates how users are changing their posting behavior pre- and post-shock. This is measured as the difference of the predicted number of posts done at the shock by the two regression models (intercept of the second model - intercept of the first model). We note that for all values of W , we observe positive discontinuity, implying a positive effect on the number of posts made by the user after receiving popularity shock. Both of these slopes are significantly different and hint towards a significant effect due to shock. Looking at the regression fits and the magnitude of discontinuity, we make the following observation:

Observation 1 (Increased Posting) *Users increase their posting behavior post shock.*

Observation 2 (Short-Term Gains) *Though users increase their posting behavior post shock, it also quickly decays off, as time progresses.*

Note that while the trend of the fit of the model pre-shock is positive and post-shock is negative - this could be due to the sensitivity towards our shock detection algorithm. Our shock detection algorithm works by binning the posts and classifying if a particular bin is a shock bin or not, and also, the algorithm takes into account total views rather than the average number of views. Therefore, users might be posting a high number of posts that were getting a sizable number of views (lesser than our threshold) until eventually tipping on the next bin and satisfying our threshold.

⁴Important to note that here, we also experimented with various values of D , θ and η and achieve similar results.

⁵We also experimented with higher order polynomial regression models, and results were consistent. Although we do observe overfitting in some cases.

Significance of Result

We perform following checks as mentioned by (Hausman and Rapson 2018). 1) We control for observable confounders to remove biases and account for variation in treatment. 2) We perform a Placebo Test to ensure no discontinuity at points where there should not be any. (Imbens and Lemieux 2008) suggests checking for any discontinuities at the median values of the running variable for the sub-samples corresponding to either side of the cut-off and using standard errors to test for no discontinuity. We do this test only for the sub-sample below the cut-off, as the points above our cut-off may have discontinuities due to potential future shocks. Say the shock occurs at the s^{th} bin from the start, then we check for any discontinuity at $\frac{s}{2}^{th}$ bin. We observe significantly less discontinuity and overlap between 99% CI intervals, implying no observable discontinuity. 3) We check for robustness of our results towards window size and polynomial order. 4) We fit regression lines without controlling for covariates and observe similar results, indicating no time-varying treatment effects.

Note that, as suggested in (Hausman and Rapson 2018), the McCrary density test (McCrary 2008) is not valid when time is the forcing variable. However, we argue that there is no manipulation in our case as users' can not preempt an imminent shock due to lack of knowledge of platform recommendation algorithm and the large magnitude of our shocks (50x more views with 1.5M difference).

Effect on Posted Content

In **RQ2**, we aim to determine if users alter the content they post after receiving a popularity shock. We characterize the content by using the posts' captions. The posts' captions can be noisy, so we take appropriate steps to develop a consistent representation from the captions. First, we preprocess the hashtags present in the caption by removing the '#' symbol from every hashtag and then use wordsegment⁶ library to segment these hashtags into separate words in order to extract their semantic meaning. Following this, we compute the similarity between the content posted in two time periods (set of bins). We represent the captions of all the posts done in that bin duration using a single feature vector and then measure their similarity. We use document embeddings to come up with the representation. We convert every post into a single vector using the document embedding of its caption. We leverage doc2vec (Le and Mikolov 2014) to generate embeddings.

Subsequently, we obtain a single vector representation for a time period by averaging the document embedding vectors corresponding to a set of posts from that temporal bin. We use cosine similarity to compare vectors formed using document embeddings. Cosine Similarity yields a score between 0 and 1, with 1 representing the same vectors. With the above experimental framework, we compare content posted in the shock bin with that of W bins just before and after the shock to capture the change around the shock. We also perform the analysis for users whose discontinuity in post-

⁶<http://www.grantjenks.com/docs/wordsegment/>

Time Period	All Users	
	Sim(Pre, Shock)	Sim(Post, Shock)
7	0.625 ± 0.22	0.714 ± 0.17***
30	0.656 ± 0.21	0.699 ± 0.20***
High Discontinuity Users		
7	0.645 ± 0.24	0.730 ± 0.20***
30	0.670 ± 0.21	0.732 ± 0.19***

Table 4: Results showing similarity of content for before and after the shock to the shock (** $p < 0.001$).

ing frequency lied in Top 25 percent. Based on the results in Table 4, we make the following observations:

Observation 3 (Post Shock Similarity) *Users, post-shock generate more similar content to the shock inducing posts.*

Observation 4 (High Discontinuity Similarity) *Users who increase their posting frequency more, also tend to stay more closer content-wise to the shock related posts*

We can observe from Table 4 that similarity of *doc2vec* embeddings between post-shock and shock is significantly higher than similarity between pre-shock and shock. We use significance test and obtain $p < 0.05$ to show that these two values over all users is significantly different.

Sustainability of Popularity

In both **RQ 3** and **RQ 4**, we try to answer the questions related to the sustainability of the popularity shock. For both of the questions, we leverage *survival analysis* (Miller 2011). Survival Analysis is a popular multivariate event history modeling technique that focuses on estimating the average hazard rate of an event under consideration at a given time and also corresponding relative strength of the effect of different factors on this hazard rate, where hazard rate can be defined by $h(t) = \frac{P(T < t + \delta | T \geq t)}{\delta t}$. Cox proportional hazard model (Cox 1972) can be used to estimate this probability and the coefficients of the regression $h(t, X) = \theta(t) \exp(\beta^T X)$ using partial likelihood, without making any assumptions about the baseline hazard rate.

Our observation period for a particular user starts from the bin where the shock occurs. We define our event of interest as the point in time post the shock where there is no difference in activity level compared to pre-shock level. Specifically, we rely on the number of views to compare post-shock and pre-shock levels. We say that the increased response due to shock has faded away if we discover B consecutive bins with the number of views less than K . We set K as the 10% of the views obtained in the shock bin. We set the value of B as 3.

RQ3: Longevity of Shock Effect

In RQ3, we study how long the effect of a popularity shock lasts. We plot in Figure 1c, the survival curve for users to demonstrate the longevity of effect on shocks. From the curve, we observe that the effect of shock dies down rather

quickly for most users. For 50% of the users the effect fades away in the first 5 days itself, while it ends for 90% of the users within 39 days of the shock. This implies that it is extremely difficult to maintain response levels observed during the shock for an extended period.

Observation 5 (Shock Longevity) *Popularity shocks are short-lived. The increased response received by users goes down to pre-shock level very quickly after the shock.*

RQ4: Sustaining Shock Effect

In RQ 4, we model the factors on which the longevity of shock effect depends as well as the effect and extent of the dependence. To do this, we build on existing survival model, and use Cox Proportional Hazards regression model (Cox 1972) to quantify the effect of different factors on survival.

Factors affecting survival: We are specifically interested in understanding what a user can do to prolong the effect of popularity shock. We hypothesize the following factors:

1. *Posting frequency:* The frequency of posting represents how eager a user is to create and post more content after the popularity shock. It can be hypothesized that high posting frequency could indicate users trying to be more active on the platform and trying to engage highly with the new audience that the user has got access to. We operationalize this by the total number of posts a user does in a bin.
2. *Similarity in Consecutive Posts:* The change or variation in the content that users post could be indicative of how versatile the user is in adapting their content to the needs of their audience. A user might have got popular due to a specific type of content and keep posting it in the hope of a similar response. However, this may lead to repetitiveness in content, and the audience might lose interest. Our analysis operationalizes this by the average cosine similarity between all posts in consecutive bins.
3. *Similarity with the shock content:* The similarity between the shock-related content and the current content is an indicator of how much the user has digressed from the content, which leads to their popularity. Viewers often start associating users with a specific type of content, and thus deviating too much from that may cause disengagement from their audience. We model this as the average cosine similarity of content posted in a bin with the shock content.

Though these are the factors that we are interested in, we also control for the following variables, which could affect the longevity of the effect.

- *Effect of feedback:* The amount of feedback received by a user on the posts user created after popularity shock is indicative of the engagement levels of the user’s audience. We measure this by introducing three variables - (a) Number of likes, (b) Number of shares, and (c) Number of comments. Since these variables are highly correlated, we only use the average number of likes in the regression model.
- *Intensity of shock:* Another factor that needs to be controlled as to what was the magnitude of the shock. Higher

Covariate	HR (St Err)	LR Chisq
Avg. Likes	0.90 (0.01)***	292.43***
Shock Intensity	1.13 (0.03)***	80.59***
Posting Frequency	0.86 (0.01)***	1047.9***
Similarity between consecutive posts	6.54 (0.03)***	2734.31 ***
Similarity of posts with shock post	0.38 (0.04)***	37.57 ***

Table 5: Dependence of Shock Effect survival on other variables using Cox Regression (***) $p < 0.001$.

the intensity of the shock, higher will be the survival chance for it.

We report the results of Cox proportional hazard regression model in Table 5.

Observation 6 (Constant Posting) *Maintaining high posting frequency helps keep retaining the long-term effect.*

Observation 7 (Similarity in Content) *Users deviating away from the content which got them to the shock have shorter survival times of shock effect, at the same time having high similarity in consecutive posts can lead to repetitiveness which again causes the survival to go down.*

Observation 8 (Engagement) *On audience side, high engagement from audience helps maintain the effect of popularity shocks.*

Discussion and Implications

Research Questions

In this paper, we focused our analysis on popularity shocks. We started with four research questions related to the effects of popularity shocks, longevity, and sustainability of the shock. Specifically, **RQ1** tries to study the effect of popularity shock on users posting frequency. From the RDiT results, we discover that users increase their posting frequency after the shock compared to before. However, as time passes, the posting frequency starts to decrease. **RQ2** is concerned with analyzing how does a user changes the content that they post after popularity shock. We find that not only do users alter their content after the shock, the post-shock content is also more similar to the content which leads to the shock, as compared to before. Thus, we conclude that popularity shocks indeed induce a behavior change in users who experience them. We are interested in understanding the longevity of the popularity shock, and hence we ask the **RQ3**. We used survival analysis to answer this question. We observe that most shocks are short-lived, i.e., the shocks reduce to 10% of their shock intensity within 5 days for 50% users. For **RQ4**, we were interested in knowing the factors that enhance the sustainability of popularity shock effects. We discover that repeatedly posting the same content as well as deviating away from the shock content cause low shock survival. Finally, high posting frequency and high response received from the users lead to more prolonged shock effect survival.

It is also worth discussing that a popularity shock or virality may not always occur in a positive connotation. Such shock can also indicate hate or networked harassment (i.e. negative attention) towards the creator (Lewis, Marwick, and Partin 2021). Similarly, increased content posting frequency can be attributed to the author apologizing, explanation, or clarifications. Such hateful phenomenons can adversely affect the mental health of the creator (Patchin and Hinduja 2012) and cause instability in the community (Bilewicz and Soral 2020). Though our work is centered only on positive popularity shocks, a potential extension to our work can be to categorize shocks into positive or negative and analyze their effect on the creator’s behavior.

Implications

Our paper provides numerous insights and observations into phenomena of popularity shocks. These insights form the basis for several implications for all three - (a) advertisers, (b) platform designers, and the (c) users.

Advertisers, or brands can adjust their marketing campaigns by understanding which users are behaving in a particular fashion that will lead to lasting popularity levels. They can also use topical information to identify if popular users identify more with their brand’s content or not.

Platforms can utilize the insights from the study to devise algorithms for their trending pages. As popularity shock is found to increase users’ engagement with the platform, enhancing attention towards dormant users can cause them to resume to increase their activity. Our content similarity results also show that such shocks can cause homogenization of content on the platform.

Users can learn the behaviors which lead to sustaining the effect of popularity shock. This can help them keep their increased engagement and benefit from the shock for a longer duration.

Threats to Validity

Like any quantitative study, our work is subject to multiple threats to validity. In this section, we attempt to list biases, data issues, and threats to the validity of our study by following the framework proposed by (Olteanu, Vieweg, and Castillo 2015). First, our work is based on a single social platform, and though it works and leverages features available on multiple social platforms, similar results do not have to hold. One possible point of differentiation could be that each platform has a different recommendation algorithm for recommending content to its users. However, the effect of recommendation algorithms on our results should be minimal since we study the effect of receiving a popularity shock by the user whereas, the recommendation algorithms primarily determines who and how big of a shock user will get. Our data can also suffer from representativeness - we use just a limited set of users who posted using a limited set of hashtags. This data representation could be significantly different from the general population on the platform. Another data issue that theoretically casts clouds on the analysis is that the number of views, likes, and comments are retrospective, i.e., they are not computed in real-time while

they are the numbers on the platform at the time of data collection. Though we believe the practical effect on our results is limited since the majority of impressions on social media posts are received soon after posting (Zhao et al. 2015). For further validation, we tracked daily view counts of 1,374 randomly sampled posts for the first 10 days after posting and found that 70% of total views were received in the first 2 days. Additionally, we did perform two analyses - regression discontinuity and survival analysis. We ensured that our data and modeling choices hold the assumptions, but there might be some unobserved confounders that we might not have considered. Finally, our statistical modeling required multiple parameters related to the operationalization of theories in sociology literature. Some of these parameters might not be capturing the factors that we intended to capture or that the theories proposed.

This work forms the basis for various future works related to popularity shocks. First of all, the work can be extended to a more generalized population and more social media platforms. Similarly, extending to different users could also open the potential to study the effect of user personality or user type on how they respond to popularity shocks. Another significant improvement in this work could be by leveraging matching techniques to match users who got popular with similar content with users who did not get popular and then record average responses. This was not possible in our current work due to multiple reasons - (a) limited data and (b) the presence of too many confounders to create a propensity model for popularity prediction.

Conclusion

We performed a large-scale analysis of the effect of popularity shocks on users. Grounded in operant conditioning and increased sense of reputation, our results confirm the extent to which popularity shock leads users to post more and modify their future content to be more similar to the content that made them famous. Similarly, on analyzing the longevity of this shock, we discovered the short-lived nature of the shocks and the effects of various posting behaviors on shock longevity. We also provide factors that users could leverage for sustaining increased engagement post-popularity shock.

References

Balcazar, F.; Hopkins, B. L.; and Suarez, Y. 1985. A critical, objective review of performance feedback. *Journal of Organizational Behavior Management* 7(3-4): 65–89.

Bandura, A. 1977. Self-efficacy: toward a unifying theory of behavioral change. *Psychological review* 84(2): 191.

Bilewicz, M.; and Soral, W. 2020. Hate speech epidemic. The dynamic effects of derogatory language on intergroup relations and political radicalization. *Political Psychology* 41: 3–33.

Burgers, C.; Eden, A.; van Engelenburg, M. D.; and Buningh, S. 2015. How feedback boosts motivation and play in a brain-training game. *Computers in Human Behavior* .

Catalbas, M. C.; Cegovnik, T.; Sodnik, J.; and Gulden, A. 2017. Driver fatigue detection based on saccadic eye movements. In *2017 10th International Conference on Electrical and Electronics Engineering (ELECO)*, 913–917. IEEE.

Chen, P.-Y.; Dhanasobhon, S.; and Smith, M. D. 2008. All reviews are not created equal: The disaggregate impact of reviews and reviewers at amazon. com. *Com (May 2008)* .

Cheng, J.; Danescu-Niculescu-Mizil, C.; and Leskovec, J. 2014. How community feedback shapes user behavior. In *ICWSM*.

Cook, T. D. 2008. “Waiting for life to arrive”: a history of the regression-discontinuity design in psychology, statistics and economics. *Journal of Econometrics* 636–654.

Cox, D. R. 1972. Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)* .

Cunha, T. O.; Weber, I.; Haddadi, H.; and Pappa, G. L. 2016. The effect of social feedback in a reddit weight loss community. In *Proceedings of the 6th International Conference on Digital Health Conference*, 99–103.

Danescu-Niculescu-Mizil, C.; Kossinets, G.; Kleinberg, J.; and Lee, L. 2009. How opinions are received by online communities: a case study on amazon. com helpfulness votes. In *WWW*.

De Veirman, M.; Cauberghe, V.; and Hudders, L. 2017. Marketing through Instagram influencers: the impact of number of followers and product divergence on brand attitude. *International journal of advertising* 36(5): 798–828.

Figueiredo, F. 2013. On the prediction of popularity of trends and hits for user generated videos. In *Proceedings of the sixth ACM international conference on Web search and data mining*, 741–746.

Figueiredo, F.; Almeida, J. M.; Benevenuto, F.; and Gummadi, K. P. 2014. Does content determine information popularity in social media? A case study of YouTube videos’ content and their popularity. In *CHI*.

Figueiredo, F.; Benevenuto, F.; and Almeida, J. M. 2011. The tube over time: characterizing popularity growth of youtube videos. In *WSDM*.

Fleiss, J. L. 1971. Measuring nominal scale agreement among many raters. *Psychological bulletin* 76(5): 378.

Freberg, K.; Graham, K.; McGaughey, K.; and Freberg, L. A. 2011. Who are the social media influencers? A study of public perceptions of personality. *Public relations review* .

Gabielkov, M.; Ramachandran, A.; Chaintreau, A.; and Legout, A. 2016. Social clicks: What and who gets read on Twitter? In *ACM SIGMETRICS ICMMCS*.

Hausman, C.; and Rapson, D. S. 2018. Regression discontinuity in time: Considerations for empirical applications. *Annual Review of Resource Economics* 10: 533–552.

Hermenda, A.; Sumarwan, U.; and Tinaprillia, N. 2019. The effect of social media influencer on brand image, self-concept, and purchase intention. *Journal of Consumer Sciences* 4(2): 76–89.

Imbens, G. W.; and Lemieux, T. 2008. Regression discontinuity designs: A guide to practice. *Journal of econometrics* 142(2): 615–635.

Jin, S. V.; Muqaddam, A.; and Ryu, E. 2019. Instafamous and social media influencer marketing. *Marketing Intelligence & Planning* .

Keegan, B.; Gergle, D.; and Contractor, N. 2012. Staying in the loop: Structure and dynamics of Wikipedia’s breaking news collaborations. In *Proceedings of the Eighth Annual International Symposium on Wikis and Open Collaboration*, 1–10.

Keegan, B.; Gergle, D.; and Contractor, N. 2013. Hot off the wiki: Structures and dynamics of Wikipedia’s coverage of breaking news events. *American behavioral scientist* .

- Keegan, B. C.; and Brubaker, J. R. 2015. 'Is' to 'Was' Coordination and Commemoration in Posthumous Activity on Wikipedia Biographies. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, 533–546.
- Kollock, P.; et al. 1999. The economies of online cooperation: Gifts and public goods in cyberspace. *Communities in cyberspace* 239.
- Kopf, S. 2020. "Rewarding Good Creators": Corporate Social Media Discourse on Monetization Schemes for Content Creators. *Social Media+ Society* .
- Lamba, H.; Malik, M. M.; and Pfeffer, J. 2015. A tempest in a teacup? analyzing firestorms on twitter. In *2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*. IEEE.
- Landis, J. R.; and Koch, G. G. 1977. The measurement of observer agreement for categorical data. *biometrics* 159–174.
- Le, Q.; and Mikolov, T. 2014. Distributed representations of sentences and documents. In *International conference on machine learning*.
- Lee, D. S.; and Lemieux, T. 2010. Regression discontinuity designs in economics. *Journal of economic literature* .
- Lewis, R.; Marwick, A. E.; and Partin, W. C. 2021. "We Dissect Stupidity and Respond to It": Response Videos and Networked Harassment on YouTube. *American Behavioral Scientist* 65(5): 735–756.
- Li, X. 2013. How does online reputation affect social media endorsements and product sales? Evidence from regression discontinuity design. In *WISE*.
- Lim, X. J.; Radzol, A.; Cheah, J.; and Wong, M. W. 2017. The impact of social media influencers on purchase intention and the mediation effect of customer attitude. *Asian Journal of Business Research* 7(2): 19–36.
- Maldeniya, D.; Budak, C.; Robert Jr, L. P.; and Romero, D. M. 2020. Herding a Deluge of Good Samaritans: How GitHub Projects Respond to Increased Attention. In *Web Conference*.
- Malik, M. M.; and Pfeffer, J. 2016. Identifying platform effects in social media data. In *Tenth International AAAI Conference on Web and Social Media*.
- Mazloom, M.; Rietveld, R.; Rudinac, S.; Worring, M.; and Van Dolen, W. 2016. Multimodal popularity prediction of brand-related social media posts. In *Proceedings of the 24th ACM international conference on Multimedia*, 197–201.
- McCrary, J. 2008. Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of econometrics* 142(2): 698–714.
- McMillen, D. L.; and Austin, J. B. 1971. Effect of positive feedback on compliance following transgression. *Psychonomic Science* .
- McParlane, P. J.; Moshfeghi, Y.; and Jose, J. M. 2014. "Nobody comes here anymore, it's too crowded"; Predicting Image Popularity on Flickr. In *Proceedings of international conference on multimedia retrieval*, 385–391.
- Miller, R. 2011. *Survival Analysis*. Wiley.
- Oktay, H.; Taylor, B. J.; and Jensen, D. D. 2010. Causal discovery in social media using quasi-experimental designs. In *Proceedings of the First Workshop on Social Media Analytics*, 1–9.
- Olteanu, A.; Vieweg, S.; and Castillo, C. 2015. What to expect when the unexpected happens: Social media communications across crises. In *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*, 994–1009.
- Patchin, J. W.; and Hinduja, S. 2012. *Cyberbullying prevention and response: Expert perspectives*. Routledge.
- Pavett, C. M. 1983. Evaluation of the impact of feedback on performance and motivation. *Human Relations* 36(7): 641–654.
- Rheingold, H. 1993. *The virtual community: Finding connection in a computerized world*. Addison-Wesley Longman Publishing Co., Inc.
- Rushton, J. P.; and Teachman, G. 1978. The effects of positive reinforcement, attributions, and punishment on model induced altruism in children. *Personality and Social Psychology Bulletin* .
- Saari, J.; and Näsänen, M. 1989. The effect of positive feedback on industrial housekeeping and accidents; a long-term study at a shipyard. *International Journal of Industrial Ergonomics* 4(3): 201–211.
- Sipos, R.; Ghosh, A.; and Joachims, T. 2014. Was this review helpful to you? It depends! Context and voting patterns in online content. In *WWW*.
- Skinner, B. 1938. The behavior of animals: An experimental analysis. *New York: Appleton-Century Crofts* .
- Smit, C. R.; Buijs, L.; van Woudenberg, T. J.; Bevelander, K. E.; and Buijzen, M. 2020. The impact of social media influencers on children's dietary behaviors. *Frontiers in psychology* 10: 2975.
- Taylor, S. J.; and Letham, B. 2018. Forecasting at scale. *The American Statistician* 72(1): 37–45.
- Thistlethwaite, D. L.; and Campbell, D. T. 1960. Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational psychology* .
- Uzunoglu, E.; and Kip, S. M. 2014. Brand communication through digital influencers: Leveraging blogger engagement. *International journal of information management* .
- Veissi, I. 2017. *Influencer Marketing on Instagram*. Yayınlanmış Lisans Tezi, Haaga-Helia University Of Applied Sciences. Ph.D. thesis, Bachelor's Thesis Degree Programme in International business.
- Wang, K.; Bansal, M.; and Frahm, J.-M. 2018. Retweet wars: Tweet popularity prediction via dynamic multimodal regression. In *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 1842–1851. IEEE.
- Weng, L.; Menczer, F.; and Ahn, Y.-Y. 2013. Virality prediction and community structure in social networks. *Scientific reports* .
- Zhang, A. F.; Livneh, D.; Budak, C.; Robert, L.; and Romero, D. 2017. Shocking the crowd: The effect of censorship shocks on chinese wikipedia. In *ICWSM*.
- Zhang, A. F.; Wang, R.; Blohm, E.; Budak, C.; Robert Jr, L. P.; and Romero, D. M. 2019. Participation of new editors after times of shock on Wikipedia. In *ICWSM*.
- Zhao, Q.; Erdogdu, M. A.; He, H. Y.; Rajaraman, A.; and Leskovec, J. 2015. Seismic: A self-exciting point process model for predicting tweet popularity. In *Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining*, 1513–1522.