Investigating the Impact of Bitcoin Mentions on YouTube Engagement

DATASCI 241: Group Project

S. Cayo, M. Dietrich, T. Gibbons, T. Johnson, T. Volz

Contents

1 Abstract

This research employed a randomized controlled trial (RCT), including a pilot study and simulation, to investigate the causal impact of Bitcoin-related comments on YouTube video engagement (views, likes, and comments). YouTube videos were randomly assigned to either a control group (no comments) or a treatment group (receiving various neutral Bitcoin-related comments, e.g., "Bitcoin's decentralized nature challenges traditional financial systems"). Initial findings suggested a potential positive effect on views and comments associated with Bitcoin-related comments. However, after accounting for multiple comparisons and outliers, the evidence was not robust enough to reject the null hypothesis that Bitcoin comment injection does not meaningfully impact video engagement. Despite challenges such as sample size limitations and potential biases, this study offers insights into the complexities of social media engagement. Future research should consider larger sample sizes, improved methods for ensuring treatment adherence, and exploring heterogeneous treatment effects for a more comprehensive understanding.

2 Introduction

How do comments discussing digital assets affect video engagement? Since its inception in 2009, Bitcoin has been a controversial topic among economists, investors, and speculators. Some think of it as a store of value, while others refer to it as a Ponzi scheme. As of 2024, YouTube holds the position of the world's second largest social media platform and the world's largest video-sharing platform (Shewale, 2024), making it an ideal place to study public engagement regarding Bitcoin.

Recent studies on social media engagement provide a crucial context for understanding the factors that may influence user interaction with social platforms. In a study presented in the Journal of Research in Interactive Marketing, researchers investigated the impact of live chat comments on video engagement (Zhang, 2024). This study measured the intensity of comments at different stages of the video. Results showed intense live chat comments at the beginning of the video decreases video engagement, whereas intense live chat comments at the end of the video increases engagement.

In 2023, a separate study presented in the Online Information Review involved researchers' investigating the effect of comments interaction from content creators on video engagement (Byun, 2023). The scope of this study measured 87,232 comments from 647 YouTube videos, and showed results where video engagement is higher when content creators reply to user comments.

While these studies offer valuable insights into engagement factors, they do not directly address the causal impact of specific terminology relating to digital assets. This study aims to fill in the gap by investigating the potential causal relationship between the specific keyword "Bitcoin" in YouTube video comments and video engagement metrics (likes, views, comments).

3 Experimental Design

3.1 Comparison of Potential Outcomes

This experiment compares the engagement metrics of YouTube videos that receive comments containing the term "Bitcoin" (treatment) against those not receiving any comment injection in the control group. The primary focus is on the causal effects of Bitcoin-related comments on video engagement.

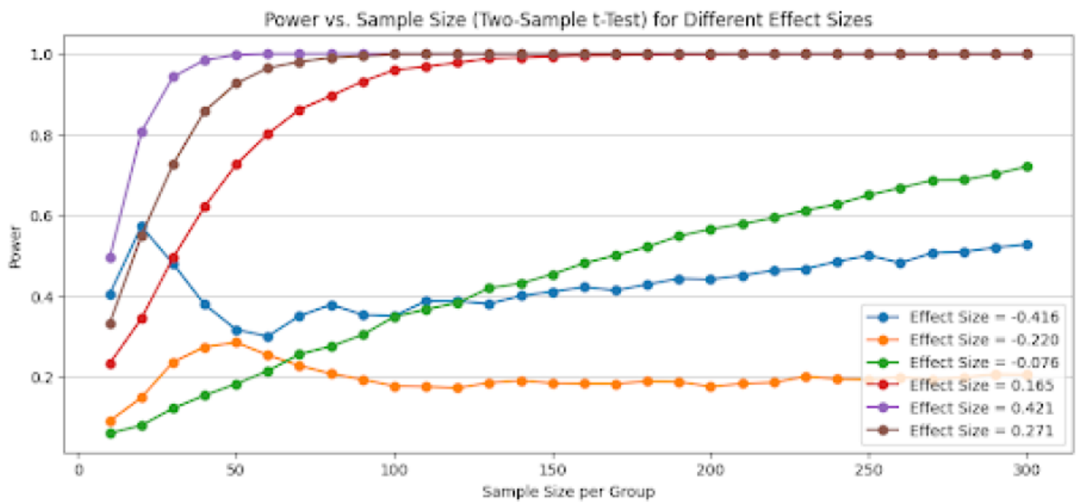
3.2 Power Analysis

We conducted a power analysis using a pilot study of 171 videos, including 53 with 'bitcoin' comments. We estimated effect sizes for views and comments, both in absolute and logarithmic formats, and calculated Cohen's d.

Our data was highly right-skewed, violating normality assumptions. We applied a log transformation, which normalized the distributions, aligning with previous research showing YouTube views follow a log-normal distribution (Cha, 2009).

- Untransformed Views: A sample size of almost 17,000 videos per group (with and without the keyword) would be required to achieve 80% power (Cohen's d = -0.034).
- Log-Transformed Data: A sample size 106 videos per group (Cohen's d = 0.40). However, it's important to interpret the results on the log scale cautiously, as the Cohen's d effect size represents relative rather than absolute differences.

3.2.1 Simulation-Based Power Analysis



To complement the analytical approach, a simulation-based power analysis was also conducted. This involved randomly generating data with and without bitcoin comments from the two distributions that we observed and comparing t-tests on the difference in the number of means.

The simulation results, as depicted in the "Power vs. Sample Size" plot, reinforce the findings from the analytical power analysis. The curves in the plot illustrate the expected behavior: statistical power increases with larger sample sizes for any given effect size.

The observed difference in means for views, on a log-mean scale, is 0.07 which our simulation shows is very hard to detect. We would need thousands of videos in our treatment and control groups to achieve a power of 0.8. The difference in the log-means of comments, however, is much larger at 1. If we measured the number of comments on a video this simulation suggests that we may only need a sample size of less than 100 to achieve a power of 0.8. We need to be cautious though, because videos with bitcoin comments may have more comments due to unobserved variables, or simply because they have more comments.

This analysis highlights that we will need to be careful about our outcome variables and we may need to perform some variable transformations to be able to achieve good statistical power.

3.2.2 Implication of Effect Size Variability

The substantial variability in the distribution of effect sizes presents a significant challenge for sample size estimation. Since the required sample size is heavily influenced by the effect size, a wide range of possible effect sizes translates to a wide range of possible sample size requirements. This uncertainty complicates the planning of a future causal experiment, making it difficult to determine a definitive sample size target.

By carefully considering the sample size, effect size, and distribution of effect size, we can increase the chances of conducting a successful experiment that will provide meaningful results. Given the variability in the data we see a range of anywhere between 100-17000 videos per group for sample size, highlighting the need to further refine YouTube video extraction and injection procedures and ensure sufficient quota to work at sufficient volume and to carefully select outcome variables that will have a large effect size.

3.3 Randomization

We conducted a randomized controlled trial using a multi-level design. To source our data, we accessed the YouTube Data API to generate a list of videos, comments, channel owner, and video titles. At first, we expanded our search to include related keywords like “cryptocurrency”, “blockchain”, “business”, “technology”, and “science” broadening the scope of our data collection. This resulted in many videos containing the word bitcoin in either the comment or title. To get a generalized list of videos, we redefined our filters for genre (“news”, “business”, “markets”, and “economy”), video-length (1-10 minutes), and language (English). This helped us maintain relevance while also circumvent the quota limitations, videos with zero views or comments were omitted.

Our data collection compiled a list of N = 393 video IDs matching our keyword search terms. We removed 36 video IDs with existing references to “Bitcoin” in the title or comment section, leaving a refined list of N = 357 eligible videos. Using simple random assignment, we allocated 188 videos to the treatment group and 169 to the control group. In the treatment group, 9 videos were excluded due to sensitive content (e.g., discussing casualties or tragic events), resulting in 179 videos receiving the treatment. Post-treatment, 4 videos in the treatment group and 3 in the control group became unavailable (deleted or made private by creators), leaving 175 and 166 videos in the treatment and control groups, respectively. Of the 175 available treatment videos, we observed attrition in 58 comments that were deleted; however, we still measured these videos so our total treatment videos still stands at 175. See figure 1

The YouTube Data API, while a valuable resource, presented certain capacity challenges that influenced our experimental design. We encountered quota limitations that restricted the number of API calls we could make within a given timeframe. This necessitated careful planning and prioritization of data retrieval tasks. Additionally, filtering videos based on specific criteria proved challenging due to the API’s inherent limitations and the vastness of YouTube’s content library. We had to iterate on our filtering strategies to balance relevance with feasibility, ensuring we collected a sufficient number of videos for our analysis while staying within the API’s constraints.

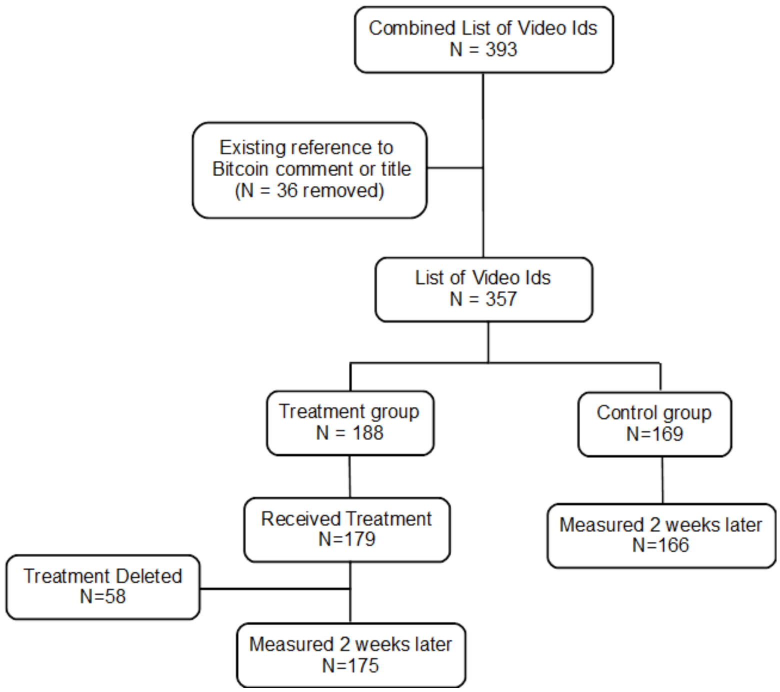
3.4 Treatment

The treatment involves injecting one of several pre-defined comments containing the term “Bitcoin” into the comment section of videos randomly assigned to the treatment group. These comments were crafted to represent a range of neutral sentiments towards Bitcoin, relating to its regulatory and technological implication while avoiding overtly positive or negative connotations that could bias viewer engagement. The control group received no comment injection. The different treatment conditions are:

- Comment #1 – “Bitcoin is an interesting development in financial technology.”
- Comment #2 – “Bitcoin seems to have its pros and cons.”
- Comment #3 – “Bitcoin’s decentralized nature challenges traditional financial systems.”
- Comment #4 – “Governments are exploring regulatory frameworks to manage Bitcoin.”
- Comment #5 – “Interesting to see how Bitcoin is used in the future.”
- Comment #6 – “Investors are closely monitoring Bitcoin’s market trends.”
- Comment #7 – “The Blockchain Technology behind Bitcoin is innovative and complex.”

The experiment was conducted over the span of one month, from the beginning of July 2024 to the beginning of August 2024. The data collection phase started in the first two weeks of July to aggregate a list of videos meeting our filter conditions and randomize videos for assignment. The treatment intervention phase took place from July 14th, 2024, to July 28th, 2024. This time frame was chosen to allow for sufficient pre-treatment data collection and reasonable time for observing potential changes in engagement metrics following comment injection. The analysis period took place after July 28th until August 5th, 2024, where we implemented a linear model to analyze pre-treatment, intervention, and post-treatment observations.

Figure 1: CONSORT Diagram



4 Data

4.1 Data Methodology

Using the YouTube Data Api, two datasets were collected to perform our analysis:

1. Video engagement data for each Video ID – each column represents information (ex: count of likes) for the video_id; each row represents a unique video_id. This dataset contained 357 rows and was used to determine attrition and experimental outcomes.

2. Comment data for each Video ID – this dataset includes a row for each extracted comment of each video_id in our randomized dataset. This dataset consisted of ~157,000 rows of data and was used to determine compliance of the treatment group. The compliance ratio was used to adjust our treatment coefficients and standard error captured in our experimental outcomes.

Using summary statistics, we confirmed the genre covariates were reasonably balanced between treatment and control. For this comparison, we confirmed the percentage difference between treatment and control was less than 5%. Numeric pre-treatment covariates (ex: pre_treatment_views) were confirmed to exhibit a high p-value greater than 0.05 for the Chi-Square analysis.

Summary Statistics

treatment		CONTROL			TREATMENT		
Variable	N	Mean	SD	N	Mean	SD	Test
genre	169			188			X2=6.928
... Blogs	13	8%		19	10%		
... Comedy	2	1%		0	0%		
... Education	17	10%		19	10%		
... Entertainment	14	8%		11	6%		
... Gaming	3	2%		5	3%		
... HowTo_Style	4	2%		5	3%		
... Music	2	1%		6	3%		
... News_Politics	89	53%		95	51%		
... Sci_Tech	23	14%		26	14%		
... Sports	2	1%		1	1%		
... Vehicles	0	0%		1	1%		
pre_treatment_views	169	99189	197389	188	115794	278307	F=0.414
pre_treatment_likes	169	2702	5212	188	3466	10939	F=0.684
pre_treatment_comments	169	663	1657	188	501	1075	F=1.222
video_length	169	312	155	188	303	155	F=0.352
comment	169			188			X2=357***
... 0	169	100%		0	0%		
... 1	0	0%		24	13%		
... 2	0	0%		28	15%		
... 3	0	0%		32	17%		
... 4	0	0%		25	13%		
... 5	0	0%		29	15%		
... 6	0	0%		26	14%		
... 7	0	0%		24	13%		
owner	169			188			X2=1.633
... Max	32	19%		42	22%		
... Sammy	39	23%		39	21%		
... Ted	41	24%		38	20%		
... Tony	28	17%		35	19%		
... Tracy	29	17%		34	18%		
post_views	166	130182	256691	184	187733	435570	F=2.206
post_likes	165	3298	6560	183	5944	27784	F=1.425
post_comment_count	166	730	1795	184	647	1326	F=0.245
pull_date	169			188			X2=0.002
... 07/27/24	137	81%		151	80%		
... 07/28/24	32	19%		37	20%		
compliance	169			188			X2=60.005***
... COMPLIER	169	100%		130	69%		
... NONCOMPLIER	0	0%		58	31%		
comment_delta	166	61	214	184	137	518	F=3.162*
view_delta	166	29351	81395	184	70092	221758	F=4.995**
like_delta	165	537	1996	183	2438	17222	F=1.987
log_view_delta	166	-Inf		184	8.7	2.8	
log_like_delta	155	4	2.5	172	4.8	2.7	F=7.661***
log_comment_delta	142	1.5	3	167	2.5	2.9	F=8.182***

Statistical significance markers: * p<0.1; ** p<0.05; *** p<0.01

4.2 Causal Relationship

To demonstrate the causal relationship between video engagement and treatment, we developed a series of linear models for each dependent variable of interest – Video Views (count), Video Comment (count), Video Like (count). Our models incorporated the covariates “genre” (as a factor) and pre-treatment results for each dependent variable. For the dependent variables in our models, we included raw values and log-transformed values because our preliminary power analysis suggested the distribution of variables were right-skewed. To account for the heteroskedasticity of our data, our models report robust standard errors.

The final consolidated stargazer table reflects one model for each dependent variable. Model selection criteria was based on the model with the lowest standard error of the treatment group. P-values captured in the final consolidated stargazer were Bonferroni-corrected to account for multiple comparisons of each dependent variable.

5 Results

The table below shows the results of linear models for the Intent to Treat and Complier Average Causal Effects (CACE) of the receiving a bitcoin comment on the change in views, comments, and likes that a video receives in the two weeks after receiving treatment. To compute the CACE we calculate a takeup rate of 69% (Of 188 videos in treatment, 58 did not have a treatment comment at the end of the study). At first glance this suggests that posting one of our bitcoin comments increases views by 46,955.4 with a 95% confidence interval of (5,854.5, 88,056.4) and increases comments by 135.1 with a 95% confidence interval of (9.121, 261.123). These are significant at the level of 0.05, but when a Bonferroni correction (3 comparisons) is applied the p-values increase to 0.078 for the effect on views and 0.108 for the effect on comments.

We also need to consider that the distributions of the change in number of views, comments, and likes are extremely right skewed so it may be possible that results are influenced due to one video on the extreme tail being assigned to treatment or control. When we inspect the data we see that the 4 videos with the largest increase in views are all in the treatment group, and the largest of these, *Vsauce’s Bubble Gum Color Science*, increased in views almost twice as much as the second largest video. Changing this one video from the treatment group to the control group increases the p-value of the change in views from 0.026 to 0.459. This video from Vsauce also has 9.94 times as many more likes in the two week period than the second most liked video and changing its assignment from treatment to control increases the p-value for additional likes from 0.107 to 0.643.

Similarly, the video with the largest increase in comments, Sky News Australia’s ‘Cringeworthy’: *Sky News host reacts to Kamala Harris being an ‘absolute clown show’* got 2.3 times more comments than the second largest increase and swapping this video from the treatment group to the control group single handedly increases the p-value for the effect on additional comments from 0.036 to 0.481. It may be possible to argue that commenting about bitcoin caused hundreds of thousands of people or bots to view Vsauce’s video, but it is more difficult to believe that the treatment caused thousands of additional comments to be posted on Sky News Australia’s video. We do not see hundreds or thousands of comments on this video about bitcoin, indicating that the bulk of the discussion is largely unrelated to our treatment. Based on our analysis of not only the regression results, but also on the distribution of the data and outliers, we do not find evidence to suggest that posting a bitcoin comment on a video produces a significant effect on the number of likes, comments, or views over the following two weeks.

Table 1: Linear Models for Views on Videos

	Dependent variable:			
	view_delta			log(view_delta)
	Assigned to Treatment	Treatment with covariates	Comment applied with covariates	Log treatment with covariates
	(1)	(2)	(3)	(4)
treatmentTREATMENT	40,741.320**	32,469.160**		0.505**
	(17,528.160)	(14,500.710)		(0.247)
	p = 0.021	p = 0.026		p = 0.042
comment1			21,827.350	
			(26,024.210)	
			p = 0.402	
comment2			10,292.150	
			(17,377.660)	
			p = 0.554	
comment3			18,306.170	
			(24,064.220)	
			p = 0.447	
comment4			23,861.050	
			(27,546.990)	
			p = 0.387	
comment5			63,092.620	
			(65,410.540)	
			p = 0.335	
comment6			73,084.630	
			(45,238.490)	
			p = 0.107	
comment7			16,055.730	
			(13,687.770)	
			p = 0.241	
pre_treatment_views		0.341**	0.341**	0.00000***
		(0.141)	(0.138)	(0.00000)
		p = 0.016	p = 0.014	p = 0.0001
Constant	29,350.550***	−457,125,233.000	−545,792,796.000	30,725.640**
	(6,316.472)	(904,709,815.000)	(882,006,914.000)	(12,753.070)
	p = 0.00001	p = 0.614	p = 0.537	p = 0.016
Observations	350	350	350	349
Residual Std. Error	170,297.500 (df = 348)	146,012.700 (df = 332)	146,359.800 (df = 326)	2.296 (df = 331)
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 1: Linear Models for CACE in Views on Videos

	Dependent variable:			
	view_delta			log(view_delta)
	Assigned to Treatment	Treatment with covariates	Comment applied with covariates	Log treatment with covariates
	(1)	(2)	(3)	(4)
treatmentTREATMENT	58,918.220**	46,955.400**		0.730**
	(25,348.410)	(20,970.250)		(0.358)
	p = 0.021	p = 0.026		p = 0.042
comment1			31,565.700	
			(37,635.010)	
			p = 0.402	
comment2			14,884.030	
			(25,130.770)	
			p = 0.554	
comment3			26,473.540	
			(34,800.560)	
			p = 0.447	
comment4			34,506.750	
			(39,837.190)	
			p = 0.387	
comment5			91,241.640	
			(94,593.700)	
			p = 0.335	
comment6			105,691.600	
			(65,421.810)	
			p = 0.107	
comment7			23,219.050	
			(19,794.620)	
			p = 0.241	
pre_treatment_views		0.493**	0.493**	0.00001***
		(0.204)	(0.200)	(0.00000)
		p = 0.016	p = 0.014	p = 0.0001
Constant	42,445.410***	−661,073,414.000	−789,300,352.000	44,434.000**
	(9,134.590)	(1,308,349,579.000)	(1,275,517,691.000)	(18,442.900)
	p = 0.00001	p = 0.614	p = 0.537	p = 0.016
Observations	350	350	350	349
Residual Std. Error	170,297.500 (df = 348)	146,012.700 (df = 332)	146,359.800 (df = 326)	2.296 (df = 331)
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 1: Linear Models for Comments on Videos

	Dependent variable:			
	comment_delta			log_comment_delta
	Assigned to Treatment	Treatment with covariates	Comment applied with covariates	Log treatment with covariates
	(1)	(2)	(3)	(4)
treatmentTREATMENT	76.826*	93.435**		1.113***
	(41.658)	(44.454)		(0.355)
	p = 0.066	p = 0.036		p = 0.002
comment1			16.779	
			(33.618)	
			p = 0.618	
comment2			−26.391	
			(29.064)	
			p = 0.364	
comment3			96.373	
			(98.035)	
			p = 0.326	
comment4			48.618	
			(53.199)	
			p = 0.361	
comment5			25.490	
			(59.121)	
			p = 0.667	
comment6			344.731	
			(232.458)	
			p = 0.139	
comment7			−18.215	
			(26.309)	
			p = 0.489	
pre_treatment_comments		0.086**		0.0001
		(0.043)		(0.0002)
		p = 0.048		p = 0.594
pre_treatment_views			0.0004*	
			(0.0002)	
			p = 0.065	
Constant	60.663***	256,469.500	−689,198.100	8,713.981
	(16.640)	(2,347,347.000)	(2,193,516.000)	(16,913.550)
	p = 0.0003	p = 0.913	p = 0.754	p = 0.607
Observations	350	350	350	309
Residual Std. Error	403.585 (df = 348)	390.741 (df = 332)	389.752 (df = 326)	2.966 (df = 291)
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 1: Linear Models for CACE in Comments on Videos

	Dependent variable:			
	comment_delta			log_comment_delta
	Assigned to Treatment	Treatment with covariates	Comment applied with covariates	Log treatment with covariates
	(1)	(2)	(3)	(4)
treatmentTREATMENT	111.103*	135.122**		1.610***
	(60.243)	(64.287)		(0.513)
	p = 0.066	p = 0.036		p = 0.002
comment1			24.265	
			(48.617)	
			p = 0.618	
comment2			−38.166	
			(42.031)	
			p = 0.364	
comment3			139.370	
			(141.773)	
			p = 0.326	
comment4			70.309	
			(76.933)	
			p = 0.361	
comment5			36.863	
			(85.498)	
			p = 0.667	
comment6			498.535	
			(336.170)	
			p = 0.139	
comment7			−26.341	
			(38.047)	
			p = 0.489	
pre_treatment_comments		0.124**		0.0002
		(0.063)		(0.0003)
		p = 0.048		p = 0.594
pre_treatment_views			0.001*	
			(0.0003)	
			p = 0.065	
Constant	87.728***	370,894.400	−996,686.500	12,601.760
	(24.064)	(3,394,624.000)	(3,172,162.000)	(24,459.600)
	p = 0.0003	p = 0.913	p = 0.754	p = 0.607
Observations	350	350	350	309
Residual Std. Error	403.585 (df = 348)	390.741 (df = 332)	389.752 (df = 326)	2.966 (df = 291)
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 1: Linear Models for Likes on Videos

	Dependent variable:			
	like_delta			log_like_delta
	Assigned to Treatment	Treatment with covariates	Comment applied with covariates	Log treatment with covariates
	(1)	(2)	(3)	(4)
treatmentTREATMENT	1,901.633 (1,282.705) p = 0.139	987.641 (612.398) p = 0.107		0.652** (0.263) p = 0.013
comment1			140.291 (790.491) p = 0.860	
comment2			836.450 (880.315) p = 0.343	
comment3			−320.571 (1,589.472) p = 0.841	
comment4			1,740.242 (1,191.388) p = 0.145	
comment5			7,596.238 (6,875.683) p = 0.270	
comment6			828.225 (1,123.729) p = 0.462	
comment7			1,461.812 (901.009) p = 0.105	
pre_treatment_likes		1.239*** (0.316) p = 0.0001		0.0001*** (0.00003) p = 0.001
pre_treatment_views			0.015 (0.013) p = 0.246	
Constant	536.624*** (155.380) p = 0.001	−53,371,753.000* (32,356,565.000) p = 0.100	−74,982,385.000 (77,867,336.000) p = 0.336	−2,796.437 (13,467.840) p = 0.836
Observations	348	348	348	327
Residual Std. Error	12,565.840 (df = 346)	6,278.521 (df = 330)	11,919.000 (df = 324)	2.352 (df = 309)
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 1: Linear Models for CACE in Likes on Videos

	Dependent variable:			
	like_delta			log_like_delta
	Assigned to Treatment	Treatment with covariates	Comment applied with covariates	Log treatment with covariates
	(1)	(2)	(3)	(4)
treatmentTREATMENT	2,750.053 (1,854.989) p = 0.139	1,428.280 (885.622) p = 0.107		0.943** (0.380) p = 0.013
comment1			202.883 (1,143.171) p = 0.860	
comment2			1,209.635 (1,273.070) p = 0.343	
comment3			−463.595 (2,298.620) p = 0.841	
comment4			2,516.657 (1,722.930) p = 0.145	
comment5			10,985.330 (9,943.295) p = 0.270	
comment6			1,197.740 (1,625.086) p = 0.462	
comment7			2,114.005 (1,302.998) p = 0.105	
pre_treatment_likes		1.791*** (0.457) p = 0.0001		0.0001*** (0.00004) p = 0.001
pre_treatment_views			0.022 (0.019) p = 0.246	
Constant	776.041*** (224.703) p = 0.001	−77,183,765.000* (46,792,571.000) p = 0.100	−108,436,064.000 (112,608,148.000) p = 0.336	−4,044.079 (19,476.570) p = 0.836
Observations	348	348	348	327
Residual Std. Error	12,565.840 (df = 346)	6,278.521 (df = 330)	11,919.000 (df = 324)	2.352 (df = 309)
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 1: Bitcoin Comments on Video Metrics

	Dependent variable:		
	Additional Views	Additional Comments	Additional Likes
	(1)	(2)	(3)
treatmentTREATMENT	32,469.160**	93.435**	987.641
	(4,048.300, 60,890.030)	(6.307, 180.564)	(−212.638, 2,187.919)
	p = 0.026	p = 0.036	p = 0.107
genreComedy	870.456	46.548	−553.899
	(−31,993.510, 33,734.420)	(−14.544, 107.641)	(−4,310.752, 3,202.953)
	p = 0.959	p = 0.136	p = 0.773
genreEducation	−12,997.020	−18.648	−580.426
	(−43,379.970, 17,385.930)	(−89.420, 52.125)	(−2,541.229, 1,380.377)
	p = 0.402	p = 0.606	p = 0.562
genreEntertainment	66,624.480	77.610	1,181.602
	(−56,744.420, 189,993.400)	(−55.669, 210.889)	(−1,873.435, 4,236.639)
	p = 0.290	p = 0.254	p = 0.449
genreGaming	−37,185.670	−42.007	−4,364.944
	(−81,848.820, 7,477.482)	(−120.307, 36.293)	(−9,897.710, 1,167.822)
	p = 0.103	p = 0.294	p = 0.123
genreHowTo_Style	−21,722.970	−45.902	287.186
	(−49,235.720, 5,789.788)	(−108.370, 16.566)	(−2,012.130, 2,586.503)
	p = 0.122	p = 0.150	p = 0.807
genreMusic	171,666.200**	32.138	−1,771.227
	(11,326.380, 332,006.000)	(−40.554, 104.831)	(−4,755.934, 1,213.480)
	p = 0.036	p = 0.387	p = 0.245
genreNews_Politics	−21,372.140	42.572	−1,066.215
	(−54,632.960, 11,888.680)	(−34.221, 119.365)	(−3,121.812, 989.381)
	p = 0.208	p = 0.278	p = 0.310
genreSci_Tech	10,666.170	13.353	−4,654.343**
	(−28,819.710, 50,152.050)	(−36.360, 63.066)	(−8,270.174, −1,038.512)
	p = 0.597	p = 0.599	p = 0.012
genreSports	4,598.838	−4.555	−1,277.351
	(−74,576.390, 83,774.060)	(−195.442, 186.332)	(−5,199.642, 2,644.939)
	p = 0.910	p = 0.963	p = 0.524
genreVehicles	−47,018.160	−127.077	548.618
	(−109,329.000, 15,292.660)	(−396.002, 141.849)	(−2,302.211, 3,399.447)
	p = 0.140	p = 0.355	p = 0.707
pre_treatment_views	0.341**		
	(0.064, 0.618)		
	p = 0.016		
pre_treatment_comments		0.086**	
		(0.001, 0.171)	
		p = 0.048	
pre_treatment_likes			1.239***
			(0.619, 1.858)
			p = 0.0001
video_posting_date	−643.316	0.361	−75.108*
	(−3,138.671, 1,852.040)	(−6.114, 6.835)	(−164.353, 14.138)
	p = 0.614	p = 0.913	p = 0.100
Constant	−457,125,233.000	256,469.500	−53,371,753.000*
	(−2,230,323,887.000, 1,316,073,421.000)	(−4,344,245.000, 4,857,184.000)	(−116,789,455.000, 10,045,950.000)
	p = 0.614	p = 0.913	p = 0.100
Observations	350	350	348
Residual Std. Error	146,012.700 (df = 332)	390.741 (df = 332)	6,278.521 (df = 330)
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 1: CACE of Bitcoin Comments on Video Metrics

	Dependent variable:		
	Additional Views	Additional Comments	Additional Likes
	(1)	(2)	(3)
treatmentTREATMENT	46,955.400** (5,854.465, 88,056.350) p = 0.026	135.122** (9.121, 261.123) p = 0.036	1,428.280 (−307.507, 3,164.068) p = 0.107
genreComedy	1,258.814 (−46,267.540, 48,785.170) p = 0.959	67.316 (−21.033, 155.665) p = 0.136	−801.024 (−6,234.010, 4,631.962) p = 0.773
genreEducation	−18,795.690 (−62,734.110, 25,142.730) p = 0.402	−26.967 (−129.315, 75.380) p = 0.606	−839.386 (−3,675.008, 1,996.237) p = 0.562
genreEntertainment	96,349.240 (−82,061.160, 274,759.600) p = 0.290	112.236 (−80.506, 304.978) p = 0.254	1,708.778 (−2,709.275, 6,126.832) p = 0.449
genreGaming	−53,776.200 (−118,366.000, 10,813.590) p = 0.103	−60.749 (−173.983, 52.486) p = 0.294	−6,312.380 (−14,313.610, 1,688.851) p = 0.123
genreHowTo_Style	−31,414.750 (−71,202.430, 8,372.924) p = 0.122	−66.381 (−156.719, 23.957) p = 0.150	415.315 (−2,909.850, 3,740.481) p = 0.807
genreMusic	248,255.700** (16,379.680, 480,131.800) p = 0.036	46.477 (−58.647, 151.601) p = 0.387	−2,561.467 (−6,877.813, 1,754.879) p = 0.245
genreNews_Politics	−30,907.400 (−79,007.670, 17,192.860) p = 0.208	61.566 (−49.489, 172.621) p = 0.278	−1,541.911 (−4,514.620, 1,430.798) p = 0.310
genreSci_Tech	15,424.920 (−41,677.740, 72,527.580) p = 0.597	19.310 (−52.583, 91.203) p = 0.599	−6,730.896** (−11,959.940, −1,501.848) p = 0.012
genreSports	6,650.627 (−107,848.900, 121,150.200) p = 0.910	−6.588 (−282.639, 269.464) p = 0.963	−1,847.247 (−7,519.483, 3,824.989) p = 0.524
genreVehicles	−67,995.490 (−158,106.500, 22,115.540) p = 0.140	−183.772 (−572.680, 205.135) p = 0.355	793.386 (−3,329.351, 4,916.124) p = 0.707
pre_treatment_views	0.493** (0.092, 0.894) p = 0.016		
pre_treatment_comments		0.124** (0.001, 0.247) p = 0.048	
pre_treatment_likes			1.791*** (0.895, 2.687) p = 0.0001
video_posting_date	−930.333 (−4,539.002, 2,678.335) p = 0.614	0.522 (−8.841, 9.885) p = 0.913	−108.617* (−237.679, 20.445) p = 0.100
Constant	−661,073,414.000 (−3,225,391,468.000, 1,903,244,639.000) p = 0.614	370,894.400 (−6,282,447.000, 7,024,236.000) p = 0.913	−77,183,765.000* (−168,895,520.000, 14,527,989.000) p = 0.100
Observations	350	350	348
Residual Std. Error	146,012.700 (df = 332)	390.741 (df = 332)	6,278.521 (df = 330)
Note:			*p<0.1; **p<0.05; ***p<0.01

6 Conclusion

Our randomized controlled trial investigated the causal impact of Bitcoin-related comments on YouTube video engagement. Initially, findings suggested a potential positive effect on views and comments. However, these results did not hold up after accounting for multiple comparisons and outliers. Consequently, we conclude that there is no robust evidence to support the claim that Bitcoin comment injection significantly influences video engagement. Despite the challenges faced, such as sample size limitations and potential biases from comment deletions, our study offers valuable insights into the complexities of social media engagement metrics. To build on this, future research should consider larger sample sizes, improved methods for ensuring treatment adherence, and the exploration of heterogeneous treatment effects. These steps would provide a more comprehensive understanding of how specific keywords in comments affect video engagement. In summary, while Bitcoin-related comments do not appear to significantly impact YouTube video engagement, the methodological lessons learned from this study are crucial for designing and implementing future experiments in digital and social media research.

7 Discussion

Our study faced certain limitations. First, the sample size, while sufficient for detecting moderate to large effects, might have limited our ability to identify subtle effects of Bitcoin comment injection. Second, despite our efforts to ensure treatment compliance, some comments might have been deleted, potentially introducing bias. Future research could address these limitations by employing a larger sample size and implementing more robust methods for ensuring treatment adherence. Additionally, exploring further techniques to compensate for heterogeneous treatment effects and conducting subgroup analyses could provide deeper insights into the nuanced relationship between Bitcoin-related comments and video engagement.

The capacity limitations of the YouTube Data API posed certain challenges to our research as well. The quota restrictions limited the number of videos we could collect and analyze, potentially affecting the generalizability of our findings. Additionally, the filtering process might have inadvertently excluded some relevant videos or introduced biases in the sample. Future research could explore alternative data collection methods or leverage more advanced API features to overcome these limitations and obtain a more comprehensive dataset.

8 References

Cha, M., Kwak, H., Rodriguez, P., Ahn, Y.-Y., & Moon, S. (2009). Analyzing the video popularity characteristics of large-scale user generated content systems. IEEE/ACM Transactions on Networking, 17(5), 1357–1370. https://doi.org/10.1109/tnet.2008.2011358

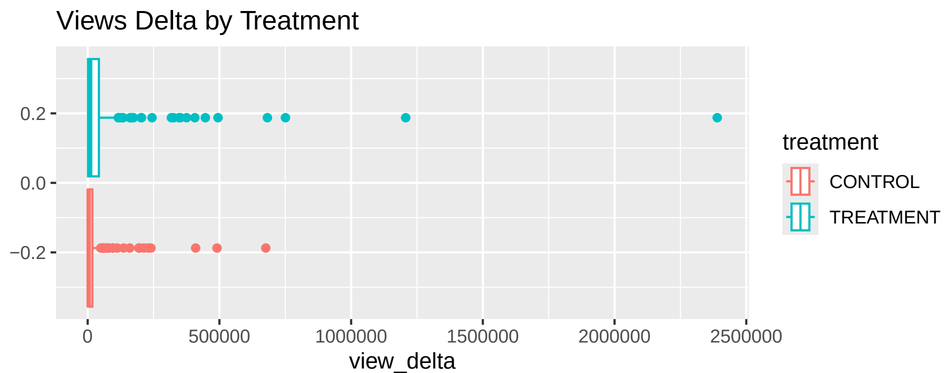
Shewale, R. (2024). YouTube statistics for 2024 (users, Facts & More). DemandSage. https://www.demandsage.com/youtube-stats/#

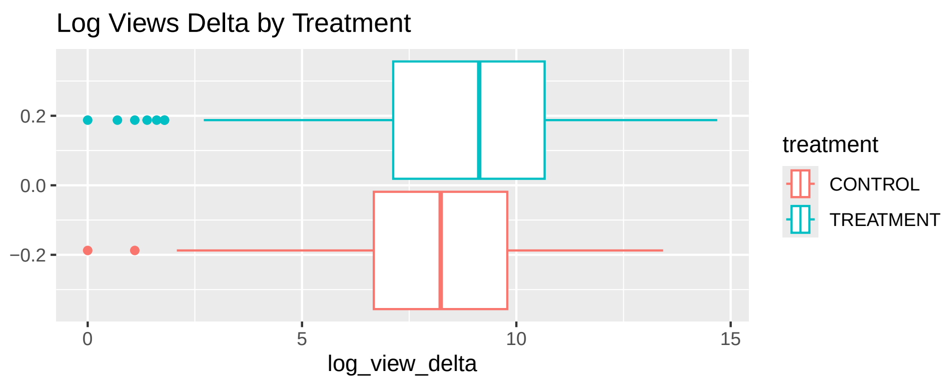
Byun, U., Jang, M. and Baek, H. (2023), "The effect of YouTube comment interaction on video engagement: focusing on interactivity centralization and creators' interactivity", Online Information Review, Vol. 47 No. 6, pp. 1083-1097. https://doi.org/10.1108/OIR-04-2022-0217

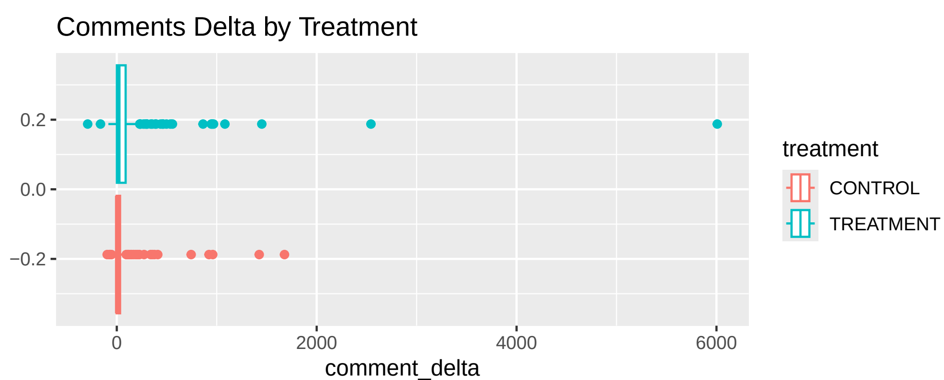
Zhang, X., Zhao, Z. and Wang, K. (2024), "The effects of live comments and advertisements on social media engagement: application to short-form online video", Journal of Research in Interactive Marketing, Vol. 18 No. 3, pp. 485-505. https://doi.org/10.1108/JRIM-02-2023-0069

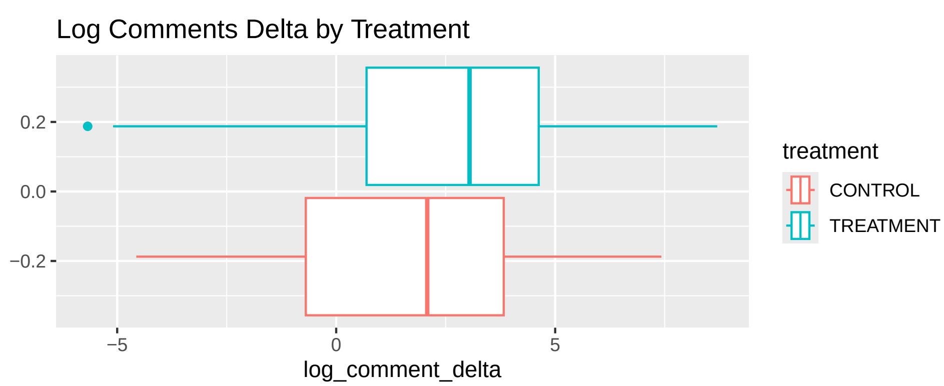
9 Appendix

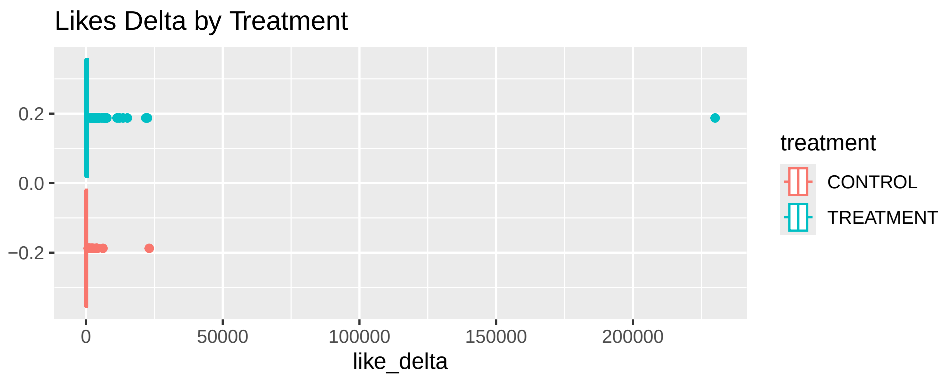
9.1 Appendix A – Visualizations of Treatment and Control

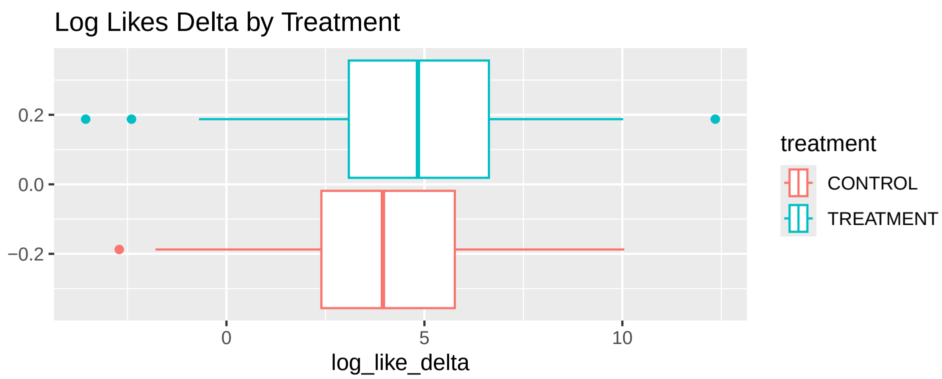




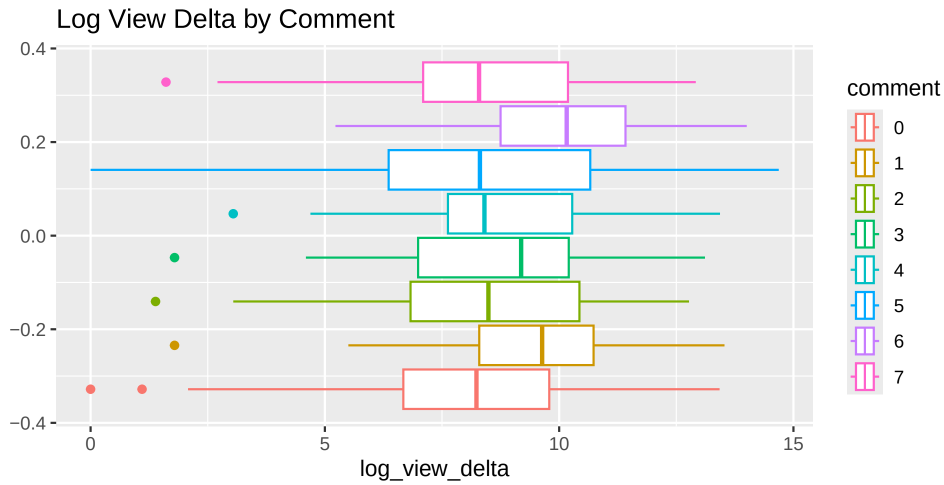


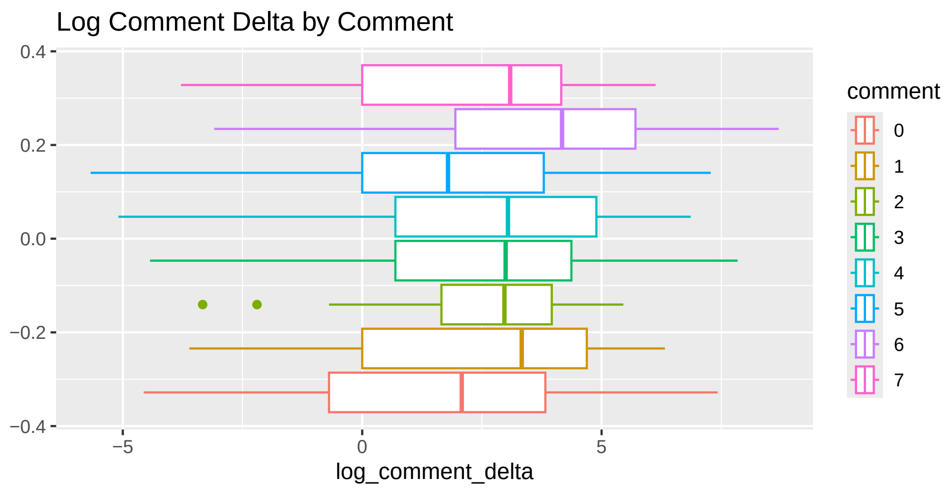


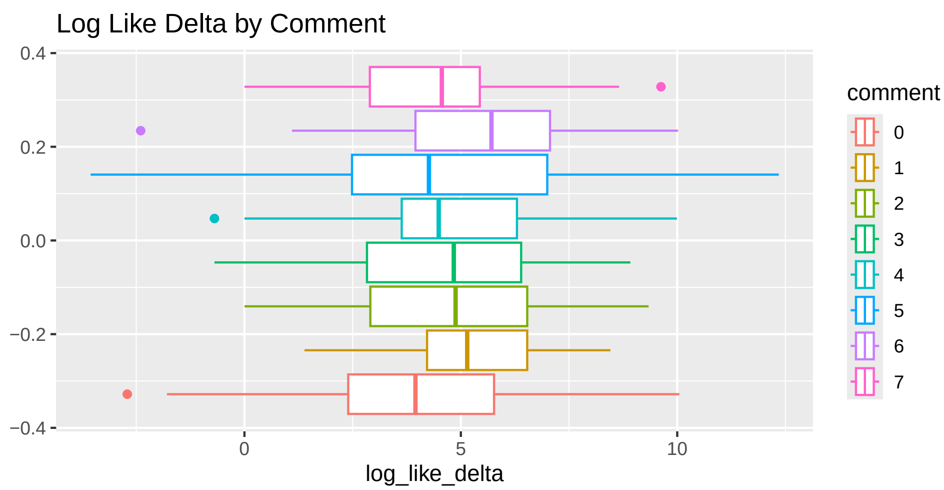

Likes Delta by Treatment
0.2
0.0
−0.2
0
50000
100000
150000
200000
like_delta
treatment
CONTROL
TREATMENT


Log Likes Delta by Treatment
0.2
0.0
−0.2
0
5
10
log_like_delta
treatment
CONTROL
TREATMENT

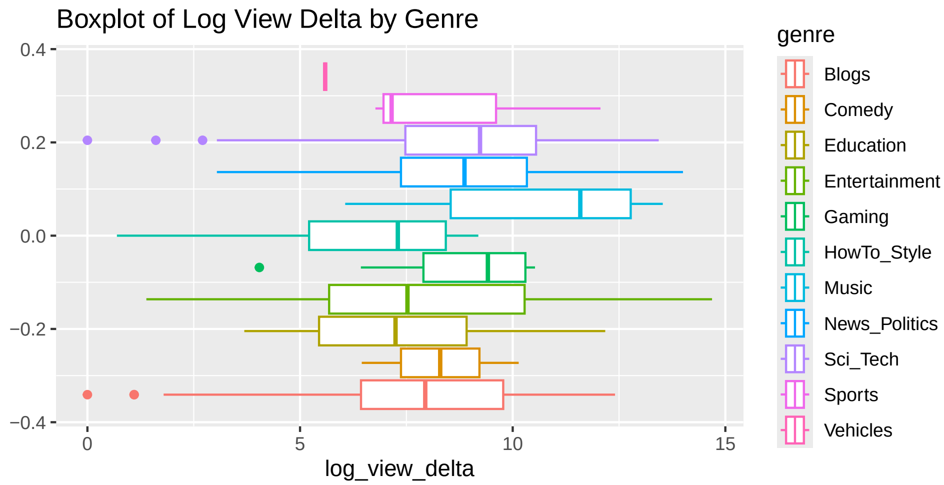
9.2 Appendix B – Visualizations of Comments

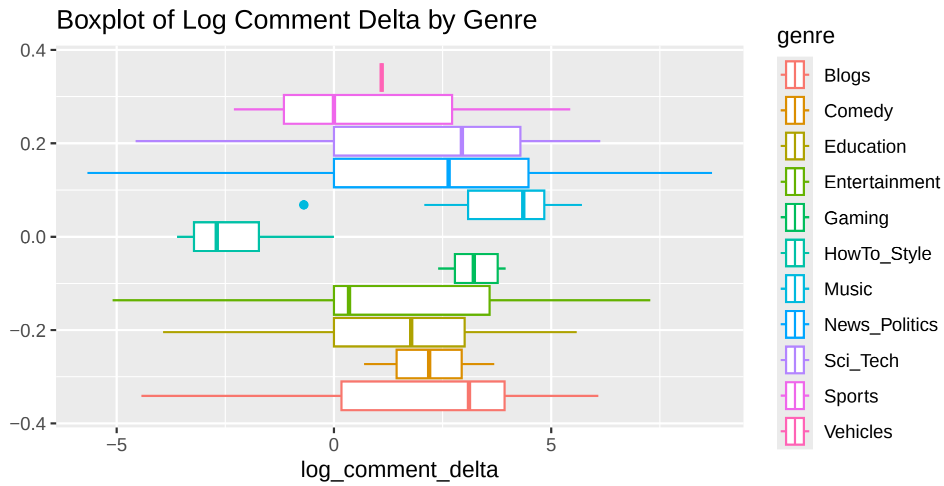


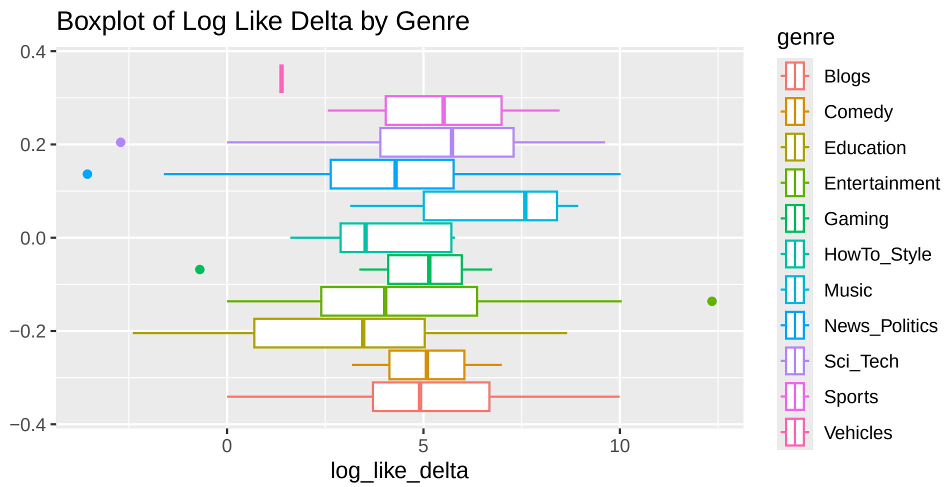




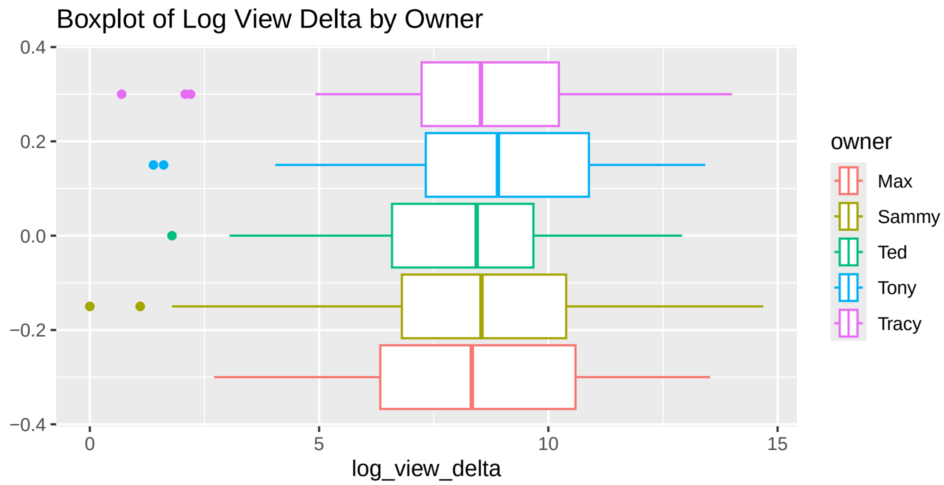
9.3 Appendix C – Visualizations of Genres

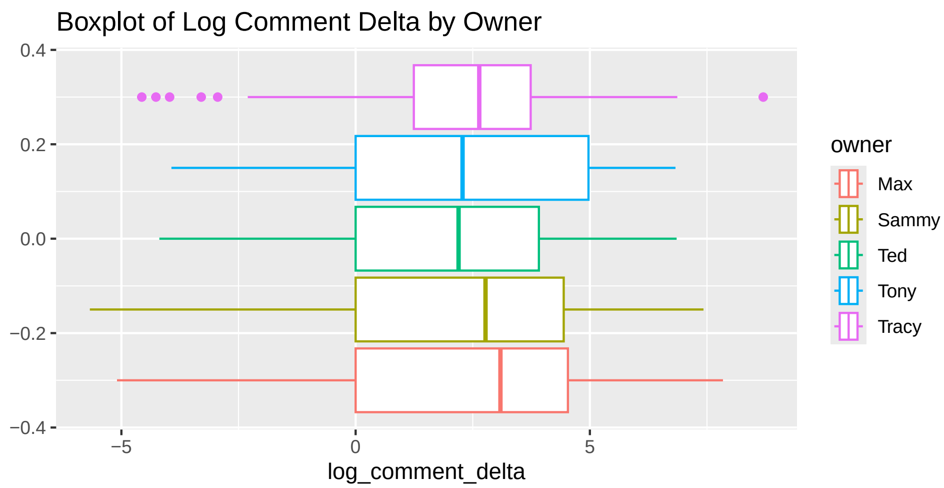


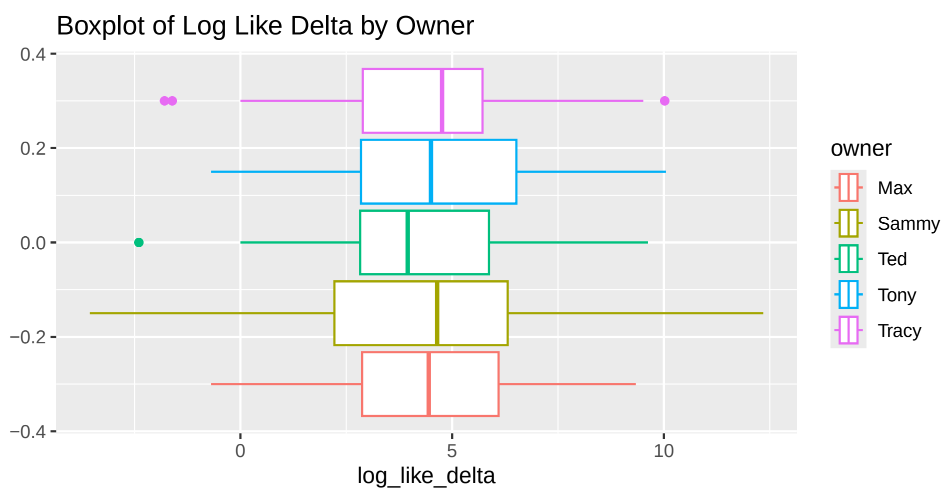




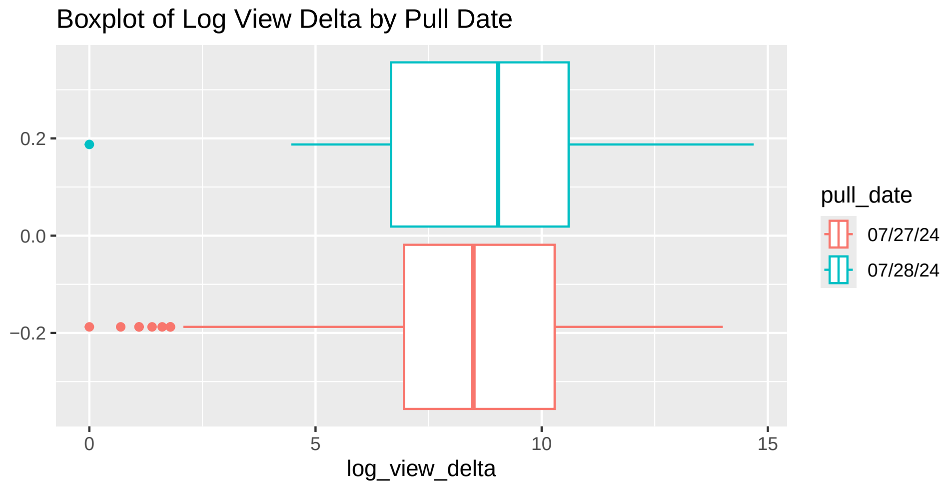
9.4 Appendix D – Visualizations of Owners

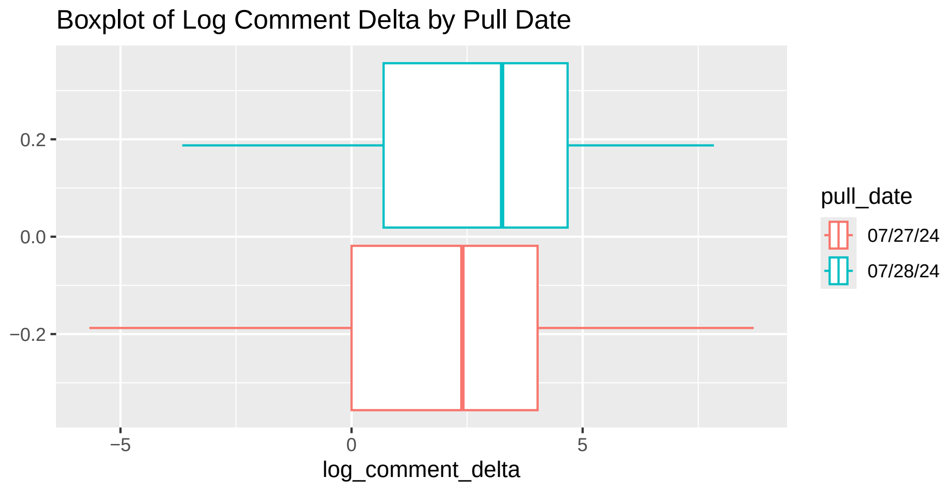


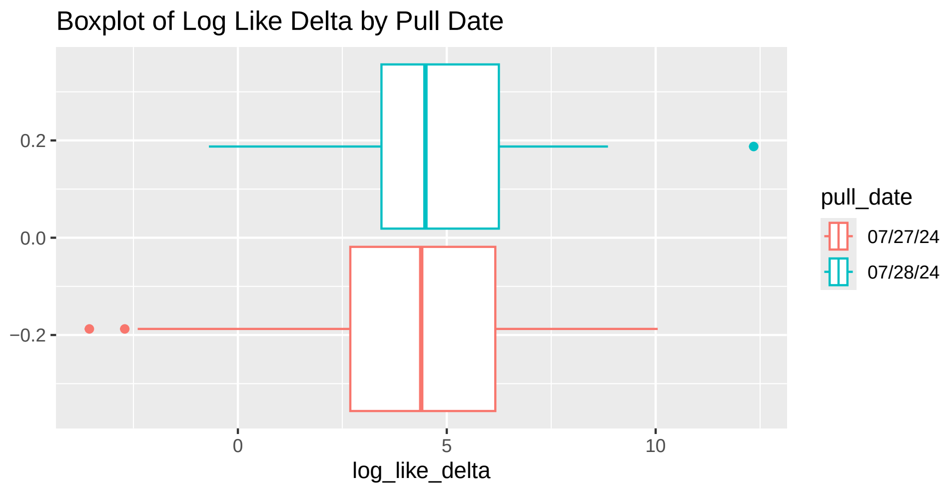




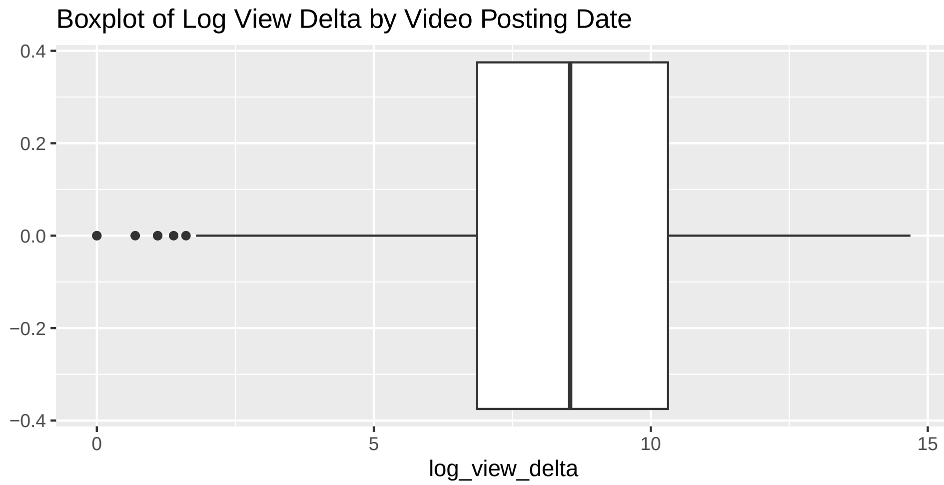
9.5 Appendix E – Visualizations of Pull Dates

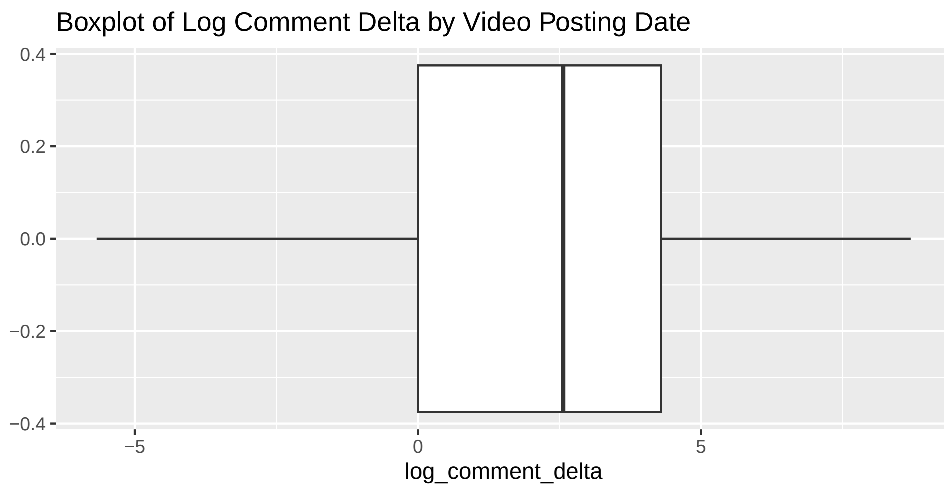


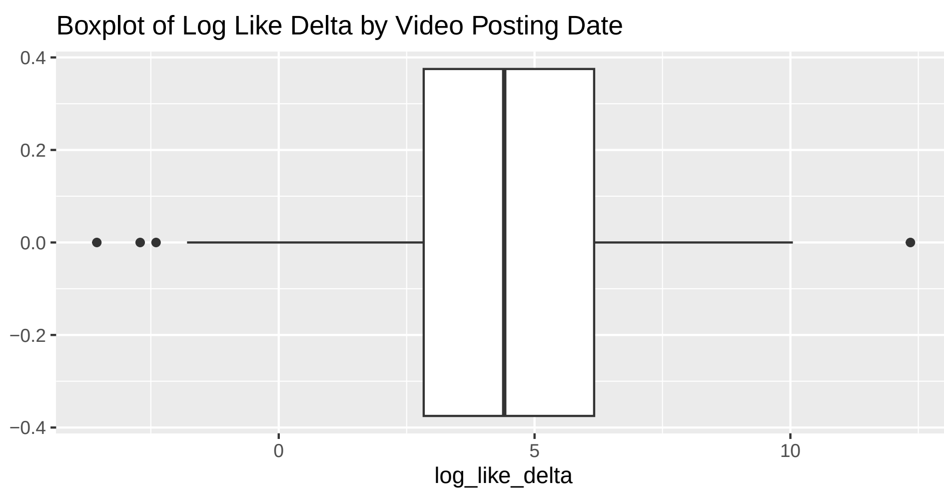




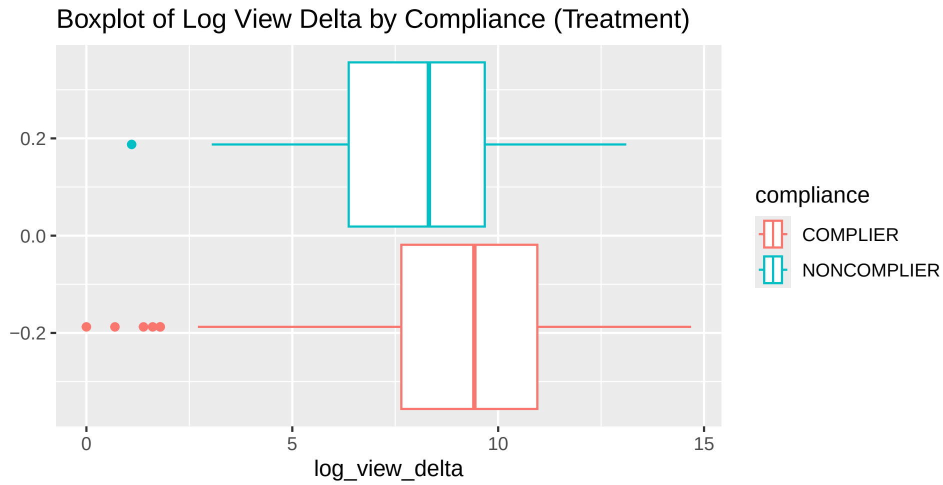
9.6 Appendix F – Visualizations of Post Dates

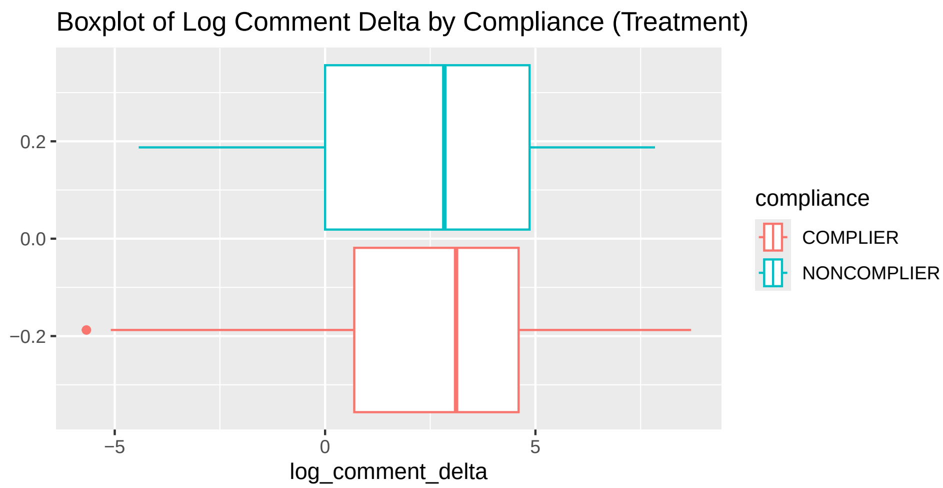


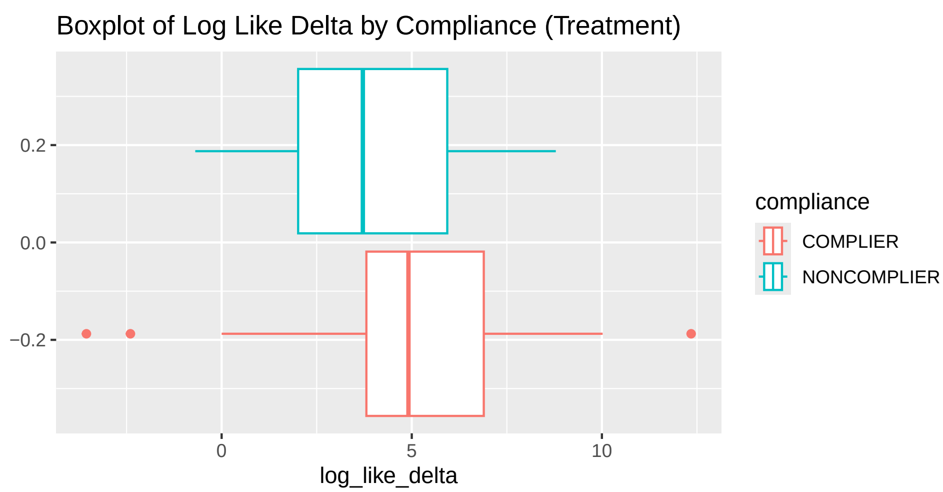




9.7 Appendix G – Visualizations of Compliance







9.8 Appendix H – Visualizations of Comments by Genre