Exploring YouTube Recommendations

Charlie Walker

1 Introduction

Recommendation systems are increasingly used to provide users with personalized content.[8] The proliferation of online content, in tandem with the growing sophistication of recommendation engines, has led to natural and important questions about the role of these recommendations in shaping consumer behaviour. Given the importance of high-quality information to a well-functioning democracy [2, 5] there has been a particular focus on the role of online recommendations on information and media markets.[12, 7, 1, 4]

This paper explores YouTube recommendations. In particular, we were interested in the types of videos YouTube recommends to its users. We focus on YouTube in part because of the recent media scrutiny on the impact of its recommendations on the polarization and accuracy of the videos its users encounter.[10]

To answer these questions we collected data on YouTube recommendations and video features. Our results are largely exploratory in nature, but we find little support for the claim that YouTube is systematically recommending polarizing content. Further, there is no evidence that YouTube is favouring content of a particular leaning. However, by defining ideological segregation using standard indices from the literature on racial segregation, we show that YouTube's recommendations are highly ideologically segregated, suggesting that concerns about "echo chambers" are well-founded.

2 Data

2.1 Recommendations

We collected data on YouTube recommendations using a web crawl. The process is illustrated in Figure 1. First we seeded our search with a politics-related search query. Queries were limited to be political in nature (and in particular, limited to U.S. political figures) to narrow the scope of the analysis. For each query we selected the top videos returned by YouTube. Then for each video we ran a fixed-depth breadth-first search of the recommendations, following the top two recommended videos at each split. Trees were truncated if a video was recommended more than once or else followed to a depth of eight.

We used sixty distinct queries for a total of five hundred search trees. A complete list of queries can be found in the Appendix. We collected information on approximately 16,000 unique videos from 5,500 channels.

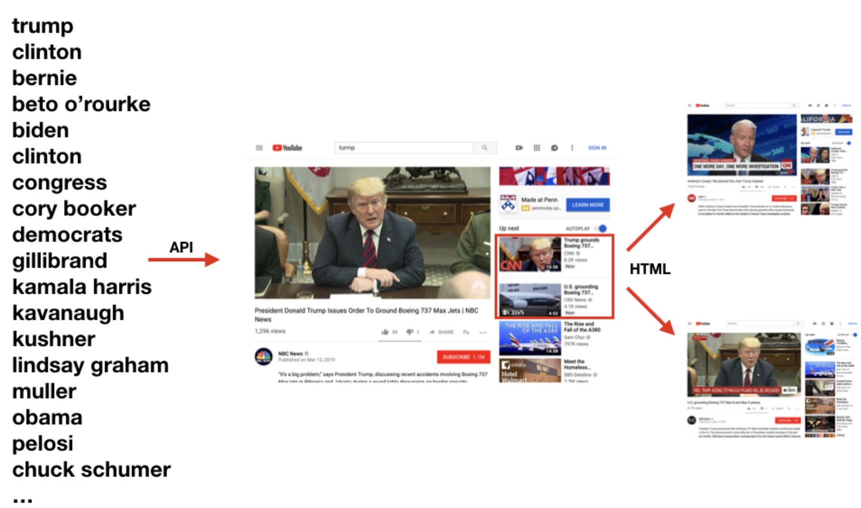

Figure 1: Crawling method. Features and recommendations were collected through a combination of API calls and HTML scraping.

We collected the following information on each video:

- title
- category
- channel
- post date
- views
- likes/dislikes
- top 20 comments and total comment count
- video description
- subtitles (where available)

Figure 12 in the Appendix illustrates the features. We minimally processed and assigned polarity scores to the text data (comments, descriptions, subtitles) using the Natural Language Toolkit (NLTK).

2.2 Content Bias

In addition to the features above we also wanted to understand political polarization. Since a precise classification of YouTube channels was not available, we used bias ratings of media outlets and matched against channel names. To capture as many channels as possible we used media bias ratings from three sources:

1. Media Bias Fact Check (`www.mediabiasfactcheck.com`): a leading classifier of the U.S. media landscape. Classifies outlets as extreme-left, left, center-left, center, center-right, right, or extreme right. We collapsed this scale to left, center, and right.
2. AdFontes (`www.adfontesmedia.com`): assigns media outlets polarity and factuality scores. Polarity is an integer rating where $polarity < 0$ indicates a left-leaning outlet, and $polarity > 0$ indicates a right-leaning outlet. For our purposes we ignored the integer scale and used the sign of the rating to classify outlets as left or right biased.
3. AllSides (`www.allsides.com`): a news website that crowdsourced bias classifications of media outlets. Bias ratings are "based on blind surveys of people across the political spectrum, multi-partisan analysis and other in-depth analyses as well as tens of thousands of user ratings." Outlets are classified as left, center-left, center, center-right, or right. As with MBFC we collapsed this scale to left, right, or center.

After matching outlets with channel names we constructed a final bias classification with a majority vote. We were able to classify approximately 200 YouTube channels, or 3% of all channels visited in our crawls. However, because some channels were visited more often than others this classification was able to account for over 20% of all unique videos in our data.

3 Results

Table 1 summarizes the main variables collected during the crawl. Text polarities are defined in the interval $[-1, 1]$, where -1 denotes negative sentiment, and $+1$ positive sentiment. The zero-valued minimums for some variables is a result of the crawl occasionally visiting "Live" videos, which do not maintain a view count. Caption, comment, and description sentiments roughly correspond on average.

Table 1: Variable Distributions

Variable	mean	median	min	max	sd
caption.polarity	0.11	0.11	-0.38	0.7	0.06
comment.polarity	0.11	0.10	-0.90	1.0	0.12
description.polarity	0.14	0.12	-1.00	1.0	0.20
dislikes	908.71	254.00	0.00	126913.0	2504.17
likes	11834.78	3637.00	0.00	788486.0	27908.86
n.comments	2462.18	826.00	0.00	130568.0	5349.15
views	1367820.57	446579.00	0.00	78727085.0	3178390.27

Table 2 shows the balance of left, center, or right-leaning videos and channels across our data. The imbalance is due to the fact that we were able to match more left-leaning channels despite the media classifications themselves being approximately balanced.

Table 2: Number and Proportion of Channel/Video Classifications

Leaning	No. Channels	Frac. Channels	No. Videos	Frac. Videos
NA	5438	0.97	12748	0.79
C	48	0.01	611	0.04
L	90	0.02	2065	0.13
R	56	0.01	678	0.04

3.1 Characteristics by Depth

First we look at how the characteristics of our videos differ by depth of the search. If YouTube was systematically pushing users towards a particular type of video we might see that reflected as we follow recommendations. Figure 2 shows average views by depth. Referring back to Table 1 we note that average views are regressing towards the sample mean.

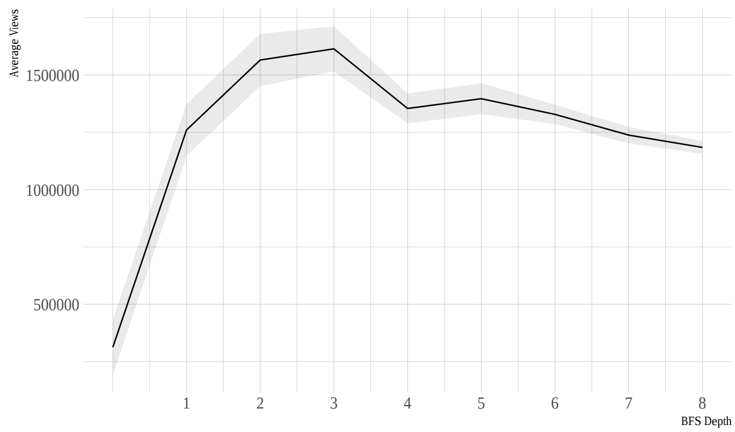

Figure 2: Average views by depth of search. Grey ribbon indicates standard error of the mean.

Next we show the average proportion of likes and dislikes as a function of tree depth. Again, we note a regression towards the mean from Table 1.

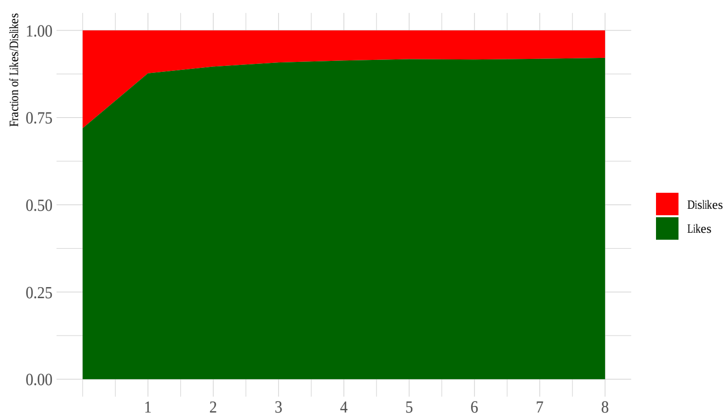

Figure 3: Average mix of likes and dislikes by tree depth. Grey ribbon indicates standard error of the mean.

One of the features were were most interested in exploring were the text sentiment scores. While there was no significant trend in the polarity of subtitles and video descriptions (see Figures 13 and 14 in the Appendix), Figure 4 shows an upwards trend in the positive valence of comments, increasing on average from 0.05 to 0.12 from the root to the leaf vertices of the recommendation trees.

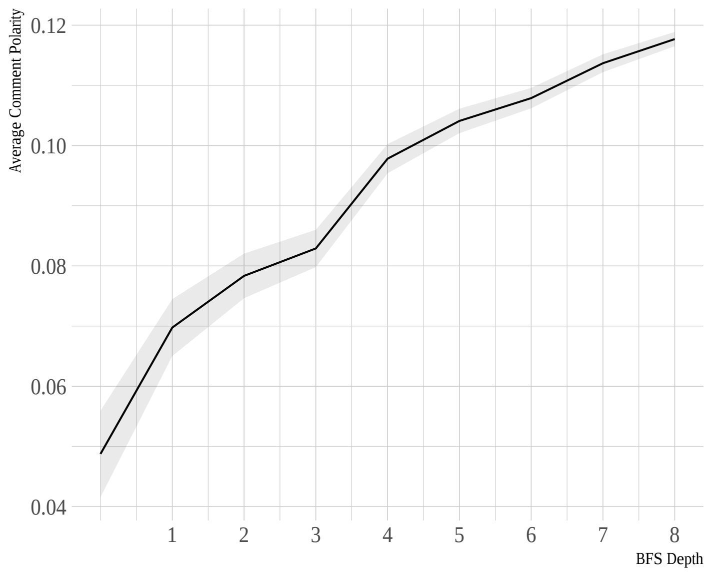

Figure 4: Average comment polarity by depth of search. Grey ribbon indicates standard error of the mean.

One potential explanation for the trends in our data is that the recommendation engine may be systematically directing users towards older. This would be a reasonable explanation for the upward trend of views by depth from Figure 2. Figure 5 shows that it is indeed the case that videos tend to get older as a function of tree depth.

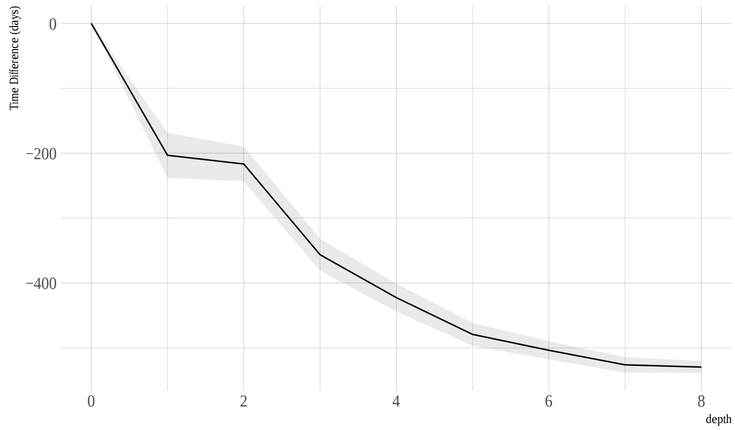

Figure 5: Relative video age by tree depth. Grey ribbon indicates standard error of the mean.

However, video age alone does not explain the trends in views and comment polarity: Figures 6 and 7 show binned averages for number of views and polarity of comments by video age, respectively.

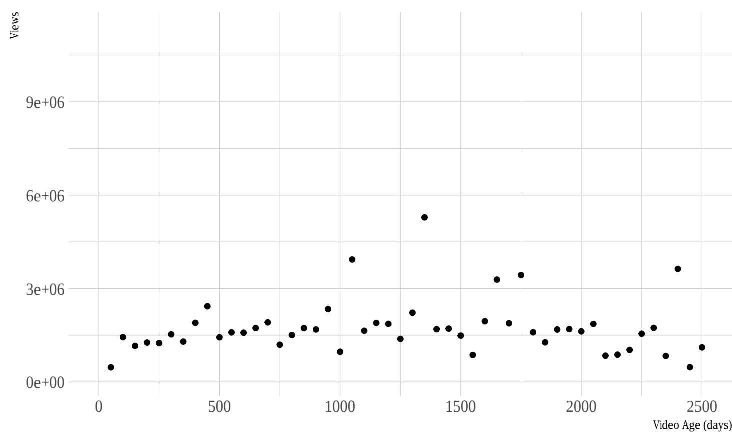

Figure 6: Average views by age of video in days. Averages were taken over 50-day bins.

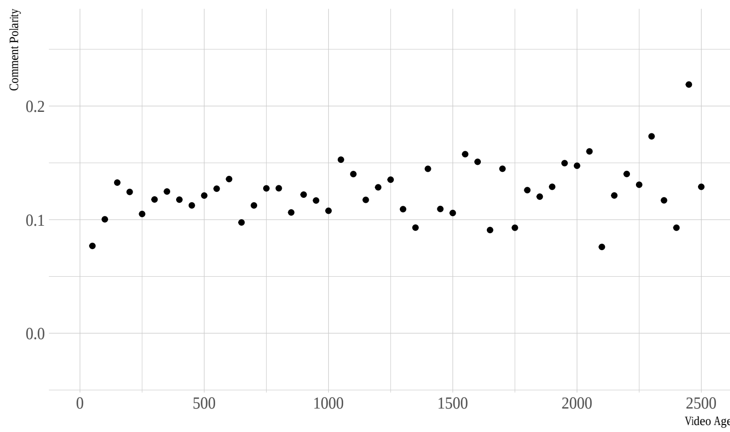

Figure 7: Average comment polarity by age of video in days. Averages were taken over 50-day bins.

Next we use the channel classifications described in Section 2.2 to see if there is any movement in the political leaning of videos with tree depth. Figure 8 suggests that this is not the case. Note that we classify videos based on the political leaning of their channel. Comparing Figure 8 to Table 6 we see that the distributions are similar regardless of depth.

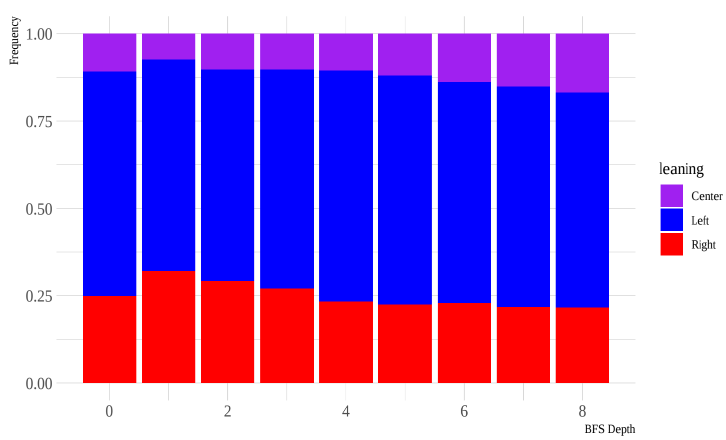

Figure 8: Political bias by tree depth.

Breaking this up by political bias of the seed video (where available) confirms that Figure 8 is reflecting the average proportions of each type of channel in our classification: regardless of the classification for the root video we ultimately revert to the mean.

Figure 9: Political bias by tree depth and bias of the seed video.

Recommendations as a Graph

Although the previous section explored average characteristics by tree depth, it is not immediately obvious that tree depth is the best dependent variable to use. It captures the process by which a user might navigate YouTube, but may not adequately capture the types of videos that YouTube is most frequently recommending, nor the "echo chamber" effect that has been the focus of substantial media coverage since the 2016 election. In this section we explore our data as a graph $G = (V, E)$: the vertex set V consists of individual videos, with directed edge $(i, j) \in E$ if video i recommends video j.

With this construction there are a number of measures we can use to understand recommendations. First we repeat the exercises from the previous section, but using vertex in-degree on the horizontal axis. Recall that, given our definition of the recommendation graph, in-degree is equivalent to the number of times our crawler visited a particular video. Table 3 shows the top ten videos in our data by in-degree: major news outlets and late night talk shows dominate. This is in contrast to other prominent work on this topic, which claimed YouTube was systematically recommending highly polarized content from fringe accounts.[11]

Table 3: Top Videos by In-Degree

Video	In-Degree	Title	Channel
jENapQxK6bc	144	Johnny Carson's Monologue Has Rough Start, But Hilarious Ending 12-14-1988	Johnny Carson
bZ83Wh0V38k	39	Funeral home sales tactics: Hidden camera investigation (Marketplace)	CBC News
h1AWLo_fK1U	36	Meet the Homeless Americans Living in Walmart Parking Lots	SBS Dateline
l41qNb2Vr-M	32	New! Ocasio-Cortez Gets SCHOOLED in Congress on The Border	Savage Nation
rs2RlZQVXBU	31	Mike Pence: Last Week Tonight with John Oliver (HBO)	LastWeekTonight
hHhYLJMi7CE	28	George Carlin- "Everyday Expressions"	doctrDave
HSW8stgUgUA	27	How President Donald Trump Hid His School Transcripts — Hardball — MSNBC	MSNBC
tQyJlKI26lk	25	Dennis Miller tells about dinner with Frank Sinatra	Joseph Stefanelli
_QBDgwtO_Xk	25	Ben Shapiro Leaves Liberal Professor SPEECHLESS In An Epic Debate	Crysta
UMhLBPPtlrY	25	Peter Attia: What if we're wrong about diabetes?	TED

Figure 10 plots a number of video characteristics against video in-degree. In contrast to the previous section there is no discernible trend for any characteristic.

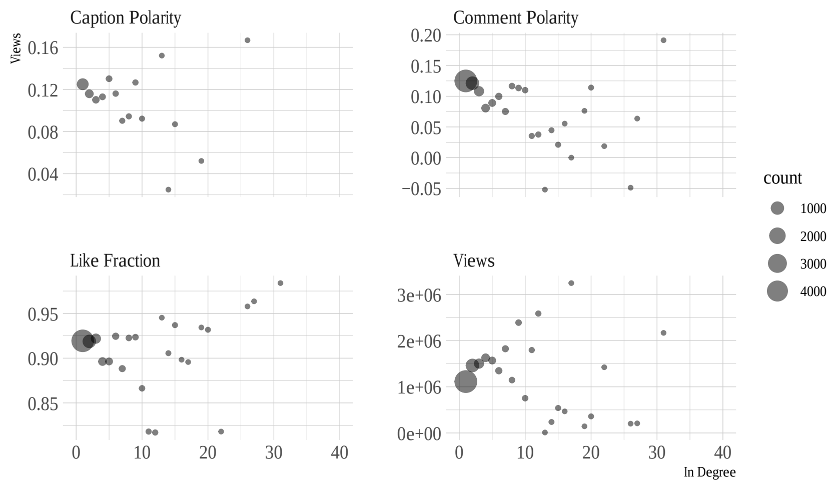

Figure 10: Video characteristics by in-degree. Size of points indicate number of videos with a given in-degree.

Of particular interest is the political leaning of recommended videos, conditional on the current video: does YouTube tend to recommend videos of a similar political leaning? Figure 11 shows the attribute mixing matrix for political leaning in the recommendation graph. The prominent diagonal entries indicate the tendency for YouTube to recommend videos of a similar political leaning, although there is still a degree of mixing between leanings.

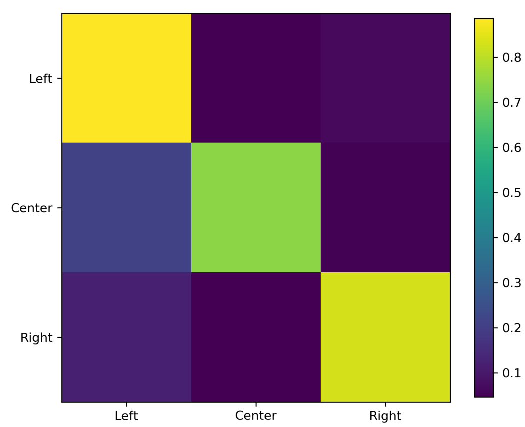

Figure 11: Mixing matrix for vertex leaning attribute. Entry (i, j) gives the probability of a video with leaning i recommending a video of leaning j.

We can formalize the strength of this association using attribute assortativity, a measure which captures the tendency for vertices in a graph to form edges with similar vertices.[9] The definition of assortativity we use is a scalar variable $\rho \in [-1, 1]$, with $\rho = -1$ indicating perfect dissortativity (vertices never form edges with similar vertices) and $\rho = 1$ indicating perfect assortativity (vertices only form edges with similar vertices). The leaning assortativity for our recommendation graph is $\rho = 0.67$. As a benchmark, the category and channel assortativities of our graph are 0.52 and 0.40, respectively.

3.2 Ideological Segregation

To get closer to the question of whether YouTube is creating echo chambers we assess the extent to which content in the recommendation graph is ideologically segregated. Our measure of segregation is the "isolation index" [3, 13], a standard metric in the study of racial segregation which has been extended to measure ideological segregation.[6].

Following [6], define R_j and L_j to be the number of conservative and liberal visits, respectively, to outlet j. Define R_m and L_m to be the total number of conservative and liberal visits on medium m, and define $n_j = R_j + L_j$ to be the total number of visitors to outlet j. The ideological segregation of users on medium m is then defined as:

$$S_m = \sum_{j \in J_m} \left(\frac{R_j}{R_m} \cdot \frac{R_j}{n_j} \right) - \sum_{j \in J_m} \left(\frac{L_j}{L_m} \cdot \frac{R_j}{n_j} \right)$$

where J_m is the set of all outlets corresponding to medium m. The fraction $\frac{R_j}{n_j}$ is the conservative share of visits to outlet j. The isolation index captures the extent to which conservatives disproportionately visit outlets whose other visitors are conservative. The index ranges from 0 (all conservative and liberal visits are to the same outlet) to 1 (all conservatives only visit 100% conservative channels and liberals only visit 100% liberal outlets). The index can be loosely interpreted as the extent to which liberals and conservatives are exposed to different facts or opinions. See [6] for details.

To measure the ideological segregation of YouTube we let j index channels. In order to use this measure we must define what we consider to be a conservative or liberal visit to a video posted by channel j. Similar to the method used to construct Figure 9, we let each crawl tree represent a "user" and classify it according to the leaning of the root video. That is, we maintain a count of conservative and liberal visits to each channel, incrementing the conservative (liberal) count of channel j by one if a tree rooted in a conservative (liberal) video visited channel j in the course of its crawl.

Table 4 shows the results and benchmarks against the measure for other mediums given by [6]. Ideological segregation on the YouTube recommendation graph is substantially higher than that of other internet media, and is similar in magnitude only to face-to-face interactions between political discussants and family.

Table 4: Ideological Segregation by Medium and Type of Interaction

	Conservative exposure of Conservatives	Liberals	Isolation Index
YouTube Recommendations	.417	.169	.247
Internet	.606	.531	.075
Offline media			
Broadcast news	.677	.660	.018
Cable	.712	.679	.033
Magazines	.587	.540	.047
Local newspapers	.695	.647	.048
Face-to-face interactions			
County	.682	.622	.059
ZIP code	.637	.543	.094
Voluntary associations	.625	.480	.145
Work	.596	.428	.168
Neighbourhood	.627	.439	.187
Family	.690	.447	.243
Political discussants	.796	.402	.394

4 Conclusion

In this paper we presented an exploratory analysis of YouTube recommendations. There is little indication that YouTube is systematically recommending polarized or fringe videos, nor is there any preference for channels of a particular political leaning. Trends in view count, comment polarity, and the like/dislike ratio do not appear to be a product of time effects alone. YouTube tends to recommend videos of a similar political leaning, but the assortativity is not perfect. We also show that the ideological segregation of YouTube recommendations are on par with that of face-to-face interactions between family and political discussants, and well above that of any other media.

We emphasize the exploratory nature of this work and do not draw any strong conclusions. There are a number of important gaps in our analysis, and we have done nothing to verify the robustness of our results. In particular, our most interesting finding - that YouTube recommendations are strongly ideologically segregated - is likely sensitive to our classification of channels, as well as our measurement strategy. At a higher level, we have collected only a small subset of the YouTube recommendation graph and do not make any claims with respect to the generalizability of our results. That said, we believe this paper is a promising proof-of-concept, and offers a number of directions for future work. The role of recommendations in media markets is still relatively unexplored, leaving room for popular dialogue to be shaped by anecdotes rather than analysis. We hope to see more work, both theoretical and empirical, on this subject in the future.

Code

Code for this project can be found at `https://github.com/cwalker4/234_project`.

References

[1] Susan Athey and Markus Mobius. “The Impact of News Aggregators on Internet News Consumption: The Case of Localization”. 2017.

[2] Gary S Becker. *Competition and Democracy.* Tech. rep. 1958, pp. 105–109.

[3] David M. Cutler, Edward L. Glaeser, and Jacob L. Vigdor. “The Rise and Decline of the American Ghetto”. In: *Journal of Political Economy* 107.3 (June 2002), pp. 455–506. ISSN: 0022-3808. DOI: 10.1086/250069. URL: https://www.journals.uchicago.edu/doi/10.1086/250069.

[4] Pranav Dandekar, Ashish Goel, and David Lee. “Biased Assimilation, Homophily and the Dynamics of Polarization”. In: (Sept. 2012). DOI: 10.1073/pnas.1217220110. URL: http://arxiv.org/abs/1209.5998%20http://dx.doi.org/10.1073/pnas.1217220110.

[5] Anthony Downs. “An Economic Theory of Political Action in a Democracy”. In: *Journal of Political Economy* 65.2 (1957), pp. 135–150. ISSN: 00223808, 1537534X. URL: http://www.jstor.org/stable/1827369.

[6] Matthew Gentzkow and Jesse M. Shapiro. “Ideological Segregation Online and Offline *”. In: *The Quarterly Journal of Economics* 126.4 (Nov. 2011), pp. 1799–1839. ISSN: 0033-5533. DOI: 10.1093/qje/qjr044. URL: https://academic.oup.com/qje/article-lookup/doi/10.1093/qje/qjr044.

[7] Ray Jiang et al. “Degenerate Feedback Loops in Recommender Systems”. In: (Feb. 2019). URL: http://arxiv.org/abs/1902.10730.

[8] Jie Lu et al. “Recommender system application developments: A survey”. In: *Decision Support Systems* 74 (June 2015), pp. 12–32. DOI: 10.1016/j.dss.2015.03.008.

[9] M. E. J. Newman. “Mixing patterns in networks”. In: *Physical Review E* 67.2 (Feb. 2003), p. 026126. ISSN: 1063-651X. DOI: 10.1103/PhysRevE.67.026126.

[10] Caroline O’Donovan et al. *We Followed YouTube’s Recommendation Algorithm Down The Rabbit Hole.* 2019. URL: https://www.buzzfeednews.com/article/carolineodonovan/down-youtubes-recommendation-rabbithole.

[11] The Guardian. *How an ex-YouTube insider investigated its secret algorithm — Technology — The Guardian.* 2018. URL: https://www.theguardian.com/technology/2018/feb/02/youtube-algorithm-election-clinton-trump-guillaume-chaslot.

[12] Marshall Van Alstyne and Erik Brynjolfsson. “Global Village or Cyber-Balkans? Modeling and Measuring the Integration of Electronic Communities”. In: *Management Science* 51.6 (June 2005), pp. 851–868. ISSN: 0025-1909. DOI: 10.1287/mnsc.1050.0363. URL: http://pubsonline.informs.org/doi/abs/10.1287/mnsc.1050.0363.

[13] Michael J. White. “Segregation and diversity measures in population distribution”. In: *Population Index* (1986), pp. 198–221.

Appendix

Tables

Queries and Classifications

Table 5: Queries and Leanings

Query	Leaning
aoc	L
ben carson	R
bernie	L
beto orourke	L
betsy devos	R
biden	L
clinton	L
cory booker	L
dan pfeiffer	L
david axelrod	L
democrats	L
dick durbin	L
eric holder	L
gavin newsom	L
gillibrand	L
james mattis	R
jeff sessions	R
jim clyburn	L
john cornyn	R
john kelly	L
john kerry	L
john lewis	L
john podesta	L
kagan	L
kamala harris	L
kavanaugh	R
kevin mccarthy	R
krysten sinema	L
kushner	R
lindsay graham	R
loretta lynch	L
maxine waters	L
mcconnell	R
mick mulvaney	R
mike pompeo	R
obama	L
paul ryan	R
pelosi	L
pence	R

rahm emanuel	L
rex tillerson	R
rick perry	R
ruth bader ginsburg	L
sarah huckabee sanders	R
schumer	L
scott pruitt	R
sotomayor	L
steve mnuchin	R
susan collins	R
susan rice	L
ted cruz	R
tim geithner	L
trump	R
wilbur ross	R
william barr	R

Channel Classifications

Table 6: Channels and Classifications

Channel	Bias
1 news	C
12 news	C
24 news	C
60 minutes australia	L
9news	L
a&e	C
aaron	L
aarp	C
abc	C
abc news	L
abc news (australia)	L
abp news	L
aj+	C
al jazeera america	C
al jazeera english	L
american conservative union	R
amv	C
ary news	R
associated press	L
at	C
bbc	L
bbc america	C
bbc news	C
ben shapiro	R
ben shapiro thug life	R

bestie	L
bfi	C
bloomberg	L
bloomberg markets and finance	C
bloomberg politics	C
breakfast club power 105.1 fm	L
business insider	C
c-span	L
cal watchdog	C
cbc	L
cbc news	R
cbn news	R
cbr	R
cbs	R
cbs news	L
cbs sunday morning	L
cbs this morning	L
cbsn	R
center for inquiry	L
cgtn	R
cgtn america	L
cjm	R
cmx	R
cnbc	C
cnbc international tv	L
cnbc television	L
cnn	L
cnn business	C
crooked media	L
ctv news	L
cut	R
daily mail	R
db	C
dbate	L
dc statesman	R
democracy now!	L
desert sky	R
dinesh d'souza	R
dsd	R
dw news	L
e4	C
eaa	C
ed	C
el pais	L
epic	L
espn	C
euractiv	L

financial times	L
forbes	R
foreign policy research institute	R
fortune magazine	R
fox	R
fox 10 phoenix	R
fox business	R
fox business	R
fox news	R
fox news insider	R
fox news shows	R
freespeechtv	L
fusion	L
glenn beck	R
gma news	L
gq	C
harvard business school	L
hbo	R
hln	R
hooverinstitution	R
hvmn	C
i.	C
ij	C
itv news	R
jason	R
jimmy kimmel live	L
jordan b peterson	R
judicial watch	R
k p	C
kansas city star	L
kgw news	C
koz	R
ko	R
krqe	C
kxly	R
kyoot	L
lastweektonight	L
late night with seth meyers	L
lbc	L
learn liberty	R
lr	C
merc	L
mg	L
mlb	L
most news	C
msnbc	L
mtv	L

national review	R
nba	L
nbc	L
nbc news	L
news24	L
newsong	L
nfb	C
nfl	C
nowthis news	L
occupy democrats	L
own	R
ozy	C
pbs	L
pbs newshour	C
pj media	R
pri	L
raw story	L
real time with bill maher	L
reasontv	R
rt	C
sal	L
saturday night live	L
sean hannity	R
sky news	L
slavb	L
snaves	L
snn	R
social justice fails	R
stars and stripes	L
tallahassee democrat	C
ted	C
tedx talks	C
the atlantic	L
the daily show with trevor noah	L
the daily wire	R
the duran	C
the economist	L
the guardian	L
the hannity	R
the late show with stephen colbert	L
the new york times	L
the new yorker	L
the oklahoman	R
the rock	L
the sun	R
the tonight show starring jimmy fallon	L
the verge	L

the view	L
the young turks	L
theellenshow	C
thegrio	L
tik	L
time	L
tlc	L
trunews	R
tv	C
ucl	R
usa news	L
usa today	C
vanity fair	L
vice	L
vice news	L
voa news	L
vox	L
wall street journal	R
washington free beacon	R
washington post	L
wfaa	C
wgbh news	L
wnyc	L
world news 7	L
wwe	C
yaftv	R
zane	R
zee news	L
zm news	C

Figures

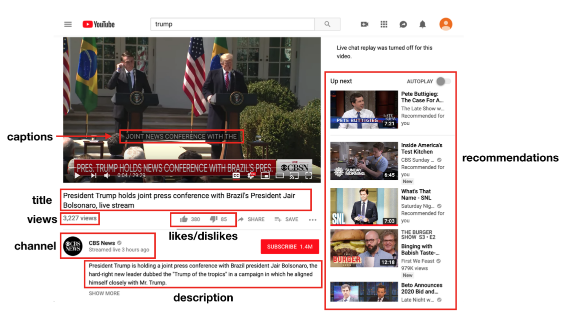

Figure 12: Sample of features collected from each video.

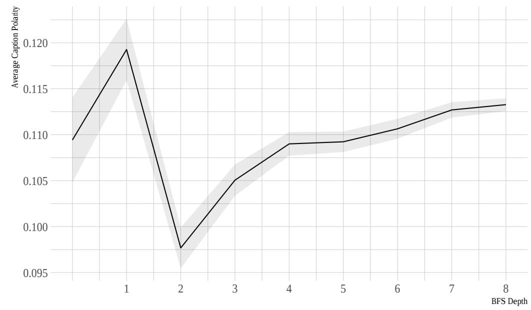

Figure 13: Subtitle polarity by tree depth. Gray ribbons indicate standard error of the mean

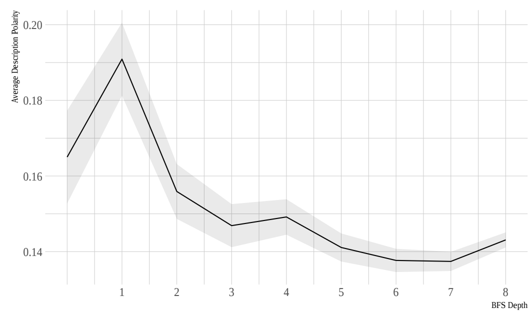

Figure 14: Description polarity by tree depth. Gray ribbons indicate standard error of the mean

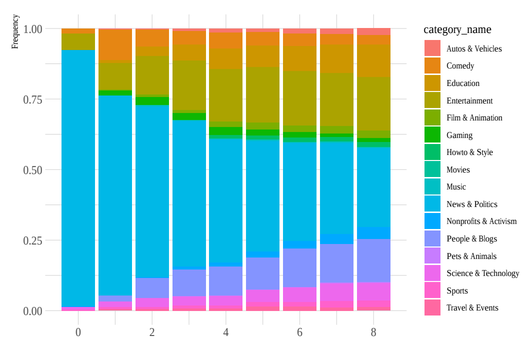

Figure 15: Category frequency by tree depth.

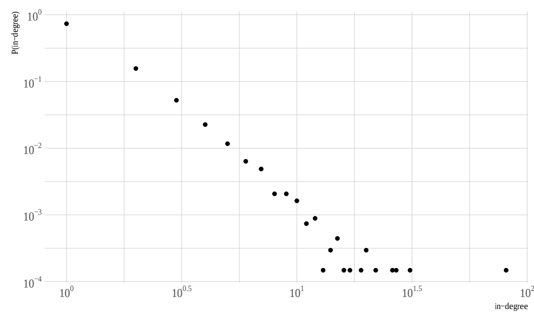

Figure 16: In-degree distribution of recommemdation graph.

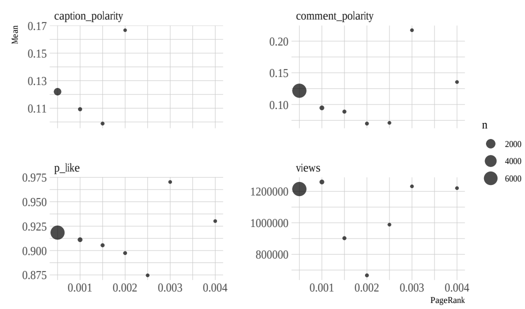

Figure 17: Video characteristics by PageRank centrality score.