Deplatforming, its dimensions and its determinants

On January 8, 2021, Twitter banned then-US President Donald Trump from using its platform [New York Times, 2021]. His account was suspended, and he was even prevented to post from his official "US President account" [BBC, 2021]. World leaders such as the German Chancellor and the Mexican President denounced Twitter's ban of Trump because it gave tech companies too much power [New York Times, 2021]. However, as the journalist Jillian C. York demonstrates, this is not something new. Online platforms such as Twitter have been engaging in "Deplatforming" Lebanese politicians, Burmese generals and even some right-wing US politicians for a long time [New York Times, 2021]. "Deplatforming" emerged as a new "form" of private censorship that is usually concerned with the internet (social media) and to lesser degree mass media [Merriam Webster, 2018; Rogers, 2020]. Due to the rising debate regarding Deplatforming, this report intends to explore people's attitudes, and their determinants, towards the different dimensions of Deplatforming through the scope of the political tolerance literature and survey experimentation. We do so by addressing three major questions**: (1) do people's attitudes towards a group influence their attitude towards the right of a company to deplatform that group**, **(2) do people view different dimensions, social, financial and technical, of Deplatforming differently or equally?**, and **(3) what determinants of political tolerance predict or explain the difference, or lack thereof, in attitudes towards Deplatforming and does the literature expectation for those determinants hold for this relatively new concept of Deplatforming?**

We attempt to understand if people's acceptance of a companies' "right" to Deplatforming a group is motivated by people's own intolerance towards that group or is it something that people accept/oppose in principle. For the purposes of this report and to apply the concepts of political tolerance in a more modern context, we follow the Sullivan et al. (1979) definition of political tolerance such that being truly tolerant is conditional on a person's support for the liberties of the group they dislike the most. Accordingly, we measure people's attitudes towards a company's right of Deplatforming of groups that they like the most and the least. We do so by fielding an online survey in which all the respondents are asked to rank six societal and/or political groups, modern adaptations of the most controversial groups in literature, from least to most liked [Boch, 2020]. We conduct an experiment by asking about half of the respondents three questions regarding a company's right of the Deplatforming of the leader of their either most or least liked group per their answers to the group ranking question. If there is a significant difference between those experimental groups, we then conclude that the difference is due to intolerance rather than a principled stand with or against Deplatforming.

To understand Deplatforming further, we explore the three experimental questions of interest. Those questions reflect the three dimensions of Deplatforming, social, financial and technical, most common in literature [Merriam Webster, 2018; Forbes, 2019; Rogers, 2020; Cato Institute, 2021]. The three questions are, in the essence, the same question in all but dimension of Deplatforming. One question asks the respondent whether they agree or disagree with a social media platform's right to deplatform a most/least liked group leader while the two other questions ask the same thing but regarding an online financial service's right and an Application/Platform hosting company's right, respectively. If there is a significant difference

among those three questions, we can then conclude that people's attitude towards Deplatforming may vary by context and thus not solely dependent on the concept of Deplatforming in and of itself.

Throughout the survey we ask several questions that are representative of the most common determinants of political tolerance [Clawson et al., 2008; Gibson et al., 2020]. Those questions cover topics such as age, gender, education, ideological leaning, political party leaning, political activity, ethnicity, religion, and degree of religiousness. We explore if those determinants extend to people's attitudes regarding the right of company to deplatform a group, both graphically and statistically through regression analysis and confidence intervals. We use the determinants of political tolerance, in literature, and not Deplatforming since there has not been an extensive analysis of Deplatforming and its determinants. This initiates a conversation regarding the report's contribution to literature and the concept of Deplatforming. We attempt to cover attitudes towards Deplatforming and its dimensions by utilizing the literature of political tolerance in a modern experimental context, and we specify the dimensions and test if they are equal. Depending on whether the dimensions are significantly equal, we may introduce an index of Deplatforming. Additionally, we include rankings of more modern groups, rather than the traditional Atheist, Communist groups utilized by older surveys [Stouffer, 1955]. We also test new questions such as a respondent's Political Activity which acts as an alternative to the more traditional Political Expertise questions [Clawson et al., 2008].

The remainder of the report is organized as follows. In Section 2, we discuss the characteristics of our survey sample in terms of its demographic and covariate distribution and

to what extent is it representative of the US population. Section 3 explores the ranked groups and people's general attitudes regarding freedom of expressions as a kind of pre-treatment analysis, before conducting the experiment and splitting the respondents into two groups. In Section 3, we solely analyze the results of the experiment and test if there is a significant effect between the two experimental groups and across the three dimensions of Deplatforming and accordingly construct the Deplatforming Index. Section 4 reanalyzes the pre-treatment analysis in Section 2 in terms of the newly constructed Deplatforming Index. Section 5 provides a graphical analysis of the determinants of interest and their effect on the Deplatforming index by the experimental treatment assignment. Section 6 utilizes a linear regression analysis to identify the significant determinants of Deplatforming. Section 7 concludes with a summary of the important results and their implications and we discuss potential extensions that may be explored in future work.

Note to the reader: All the Deplatforming questions and analyses measure a respondent's attitude towards a company's "right" to ban a group or its leader. Accordingly, throughout this report I intend to use the "acceptance" of Deplatforming to indicate that a respondent approves or agrees that a company has the "right" to ban the group or leader in question. I selected the word accept since it indicates approval but not explicit support since the question asked in the survey is framed by asking to what extent does a respondent agree or with a company's right to ban a group's leader. People may acknowledge the company's right but that does not explicitly mean that they support the Deplatforming of that group. This report uses this wording and is analyzing a company's right to deplatform a group and not whether they should deplatform the group.

Basic Demographics and Determinants (Covariates)

Variable/Covariate								
Age	Under 18 0.145%	18 – 24 3.49%	25 – 34 42.3%	35 – 44 30.6%	45 – 54 13.4%	55 – 64 7.7%	65 – 74 2.18%	75 – 84 0.145%
Gender	Male 62.5%	Female 37.5%						
Education	Less than high school 0.145%	High school graduate 7.27%	Some college 11.8%	2-year degree 7.85%	4-year degree 57%	Professional degree 14.1%	Doctorate 1.89%	
Liberal /conservative measure	Extremely liberal 16.3%	Moderately liberal 23.7%	Slightly liberal 14.8%	Neither liberal nor conservative 11%	Slightly conservative 11%	Moderately conservative 13.7%	Extremely conservative 9.44%	
Party Leaning	Republican 24.1%	Democrat 56%	Independent 18.8%	Something else 1.16%				
Political Activity in the past year*	0 8.28%	1 26.7%	2 19.9%	3 21.2%	4 23.8%			
Ethnicity	White, Non-Hispanic 76.6%	African American 8.87%	Latino or Hispanic 4.07%	Asian 6.1%	Native American 1.89%	Two or more 1.6%	Other/Prefer not to say 0.87%	
Religion	Buddhism 1.16%	Christianity 64.4%	Hinduism 1.01%	Islam 0.87%	Judaism 1.02%	None 28.3%	Other 3.2%	
Degree of religiousness	Not religious 32.8%	Slightly religious 18%	Moderately religious 23.8%	Very religious 25.3%				

Table 1. Percentage of respondents across demographic variables and variables of interest. The highest percentage of respondents in each variable is colored in red while the lowest is highlighted in blue.

The original sample contains 699 American respondents who answered the survey online on Amazon MTurk. Eleven respondents were removed from our analysis because they did not complete the entire survey and our questions of interest. The final sample utilized for this analysis includes 688 respondents. We explore the demographic distribution as well as the distribution of our determinants of interest. We start by analyzing some demographic variables like age, gender, and education. Since this is an online survey, we expect our respondents to be disproportionately young. This is exactly what we observe in this sample. The largest age group is the 25-34 group with 42.3% of the respondents belonging to that group. We observe that the percentage of respondents decreases as we move to higher age groups for every age group after the 25-34. Surprisingly, only about 3.6% of the sample is under the age of 25. The high proportion of younger respondents will likely provide a generally unrepresentative view of the American public, since young people are more involved online. Accordingly, younger people may be more biased toward something that occurs exclusively in the digital space, such as Deplatforming, than older people who rarely use the internet. The distribution of gender is also disproportionate such that 62.5%, the majority, of the respondents are Males and 37.5% are females. Thirdly, the majority of the respondent have a 4-year degree (57%) while only one respondent does not have a high school education. Those with professional degrees come in second place (14.1%) while Ph. D holders constitute the smallest group (1.89%). This makes sense since most of our respondents are relatively young and thus probably have at least a high school degree.

Covariate measures or variables of interest that are important in the literature of tolerance and censorship include a respondent's conservative/liberal ideology, party leaning, political

activity, ethnicity, religion, and their degree of religiousness. While we may observe that there is a more balanced spread in the liberal/conservative measure compared to the demographic variables. It should be noted, however, that a majority of respondents identify themselves as at least slightly liberal. While about 55% of the respondents identify as at least slightly liberal, about 34% identify as at least slightly conservative. 11% do not identify with either ideologies. While it is not a must that a conservative leans Republican and a liberal leans Democrat, party leaning seems to somehow mimic liberal/conservative leaning such that 56% of the respondents lean Democrat while about 24% lean Republican. We notice that the percentage of people who are at least slightly liberal is approximately equal to the percentage of people who lean Democrat in the sample; however, there is a Republican leaning deficit with only 24% of the respondents leaning Republican while 34% leaning at least slightly conservative. This skew may influence our results since liberals/democrats may be expected to be more in favor of Deplatforming than conservatives/republicans. It will be interesting to see if the ideological-party connection holds or will there be conflicting opinions between parties and ideologies that are traditionally aligned.

Respondents' ethnicity, religion and degree of religiousness also constitute important determinants of Deplatforming and thus require a thorough analysis. A vast majority of our respondents are Non-Hispanic White (76.6%) with about 9%, 6% and 4% of the respondents are African American, Asian and Latino or Hispanic, respectively. This percentage of white respondents may be representative of the general US population; however, there is an underestimation of the number of African Americans and Latino or Hispanic in the sample. On the contrary, the distribution of respondents by religion is very representative of religion in the

US with a majority being Christian of any denomination (65%). Only about 4% of our respondents adhere to one of the other religions while about 28% are not affiliated with any religion. The respondents' degree of religiousness is rather more balanced. About 33% of the respondents are not religious while about 25% of them are extremely religious. The rest of the respondents are in between the two extremes.

This initial analysis of demographics and covariates gives as an idea of what to expect and the degree of representativeness of the sample. We may expect that the majority of the sample would be more accepting of Deplatforming as the majority lean democratic/liberal. Conversely, we may expect stronger opinions on the topic since the sample is mostly younger respondents who may spend a long time online and thus may have been deplatformed themselves or at least witnessed it happen to someone they know. Finally, we observe that about 48% of the sample is white males, 29% is white females, 5.4% is African American males while all the other ethnic/gender categories amount to less than 5% each. This means that the attitudes observed are mostly a product of the views of those respondents. Accordingly, we now observe if these expectations and differences in respondents truly reflect a difference in attitudes towards Deplatforming.

Pre-Treatment Analysis

We asked the respondents two questions that are pivotal to our experiment. First, we asked them to rank 6 groups according to their preference from the least to most liked. Either the group they selected to be their most liked (with rank equal to 6) or the group they selected to be their least liked (with rank equal to 1) will be the group of interest that we ask them Deplatforming questions about depending on which experimental group they are assigned to.

All the six groups can be considered controversial one way or another in the US. The ranking depends on the respondents' attitude towards those groups.

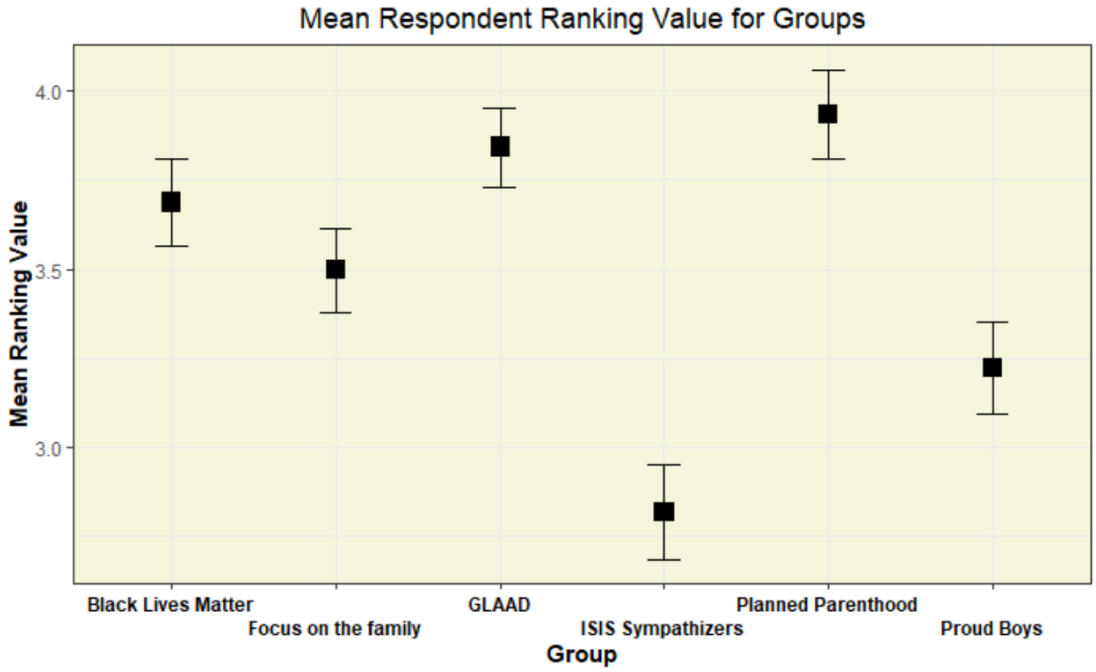


Figure 1. The respondent rankings for different groups are represented in this plot by the mean respondent ranking value for each of the 6 groups of interest. The higher the value, the more favorable a group is perceived by the respondents on average.

From Figure 1, we observe that on average, Planned Parenthood and GLAAD rank the highest among respondents. It should be noted that both Planned Parenthood, the largest provider of reproductive health services in the US, including abortions, and GLAAD, an organization that fights discrimination against LGBTQ individuals in the media to promote LGBTQ acceptance, are both considered left-leaning groups. The mean rankings between those groups are not significantly different as we observe from their overlapping confidence intervals. It is interesting and not surprising that two left-leaning groups are the top ranked groups since, as we earlier observed, the majority of our respondents identify as liberal or Democratic.

They are followed by Black Lives Matter and a bit below it is the Focus on the Family organization. Black Lives Matter, like GLAAD and Planned Parenthood, is considered a liberal leaning movement that seeks racial justice by, among other things, defunding police while Focus on the Family is a conservative leaning, fundamentalist Christian group opposing abortion and LGBTQ rights, including marriage. It is interesting that both those groups are considered statistically equal even though they may be considered on the opposite end of the spectrum. The lowest ranked group is the ISIS Sympathizers group. It is neither right or left leaning and comprises of people in the US who support the Islamic terrorist organization, ISIS, and its activities. The Proud Boys, an extreme right-wing group preaching white nationalism, is ranked higher; however, overall, fifth. It seems that, on average, our respondents rank extreme groups that support terrorism or supremacy much lower than groups that they may ideologically disagree with but are focused on social issues such as racial and sexual discrimination, promoting family values, and providing health services, including abortions.

The second pre-analysis question that is of interest to us is whether the respondents think people should be harassed or punished for their opinions in general. This question is important since it gauges the attitudes of the respondents toward freedom of expression before receiving the treatments. We first analyze the question in a vacuum and then compare the answers for this question with the respondents' attitudes regarding Deplatforming the experiment.

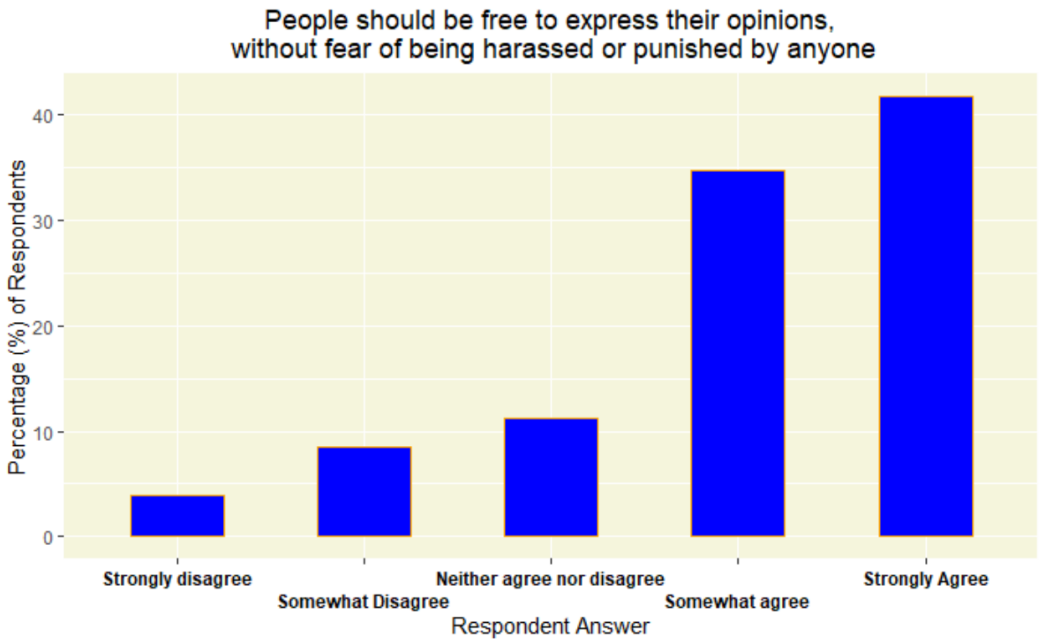


Figure 2. The percentage of respondents according to their attitude regarding if people should be free to express their opinion without fear of being punished. The question is measured using a 5-score Likert scale. The bars represent the percentage of respondents for the 5 possible answers.

From Figure 2, we observe that about 40% of the respondents strongly agree with the sentiment that people should voice their opinions freely, followed by 35% who somewhat agree. Overall, about 75% of our respondents agree to some extent with that sentiment while about 12% neither agree nor disagree and only about 13% overall at least somewhat disagree. We move to the experiment to see if these numbers or sentiments are reciprocated or not towards Deplatforming.

Treatment Analysis

As stated earlier, we attempt to analyze three aspects of Deplatforming. The dimensions include social media Deplatforming, online financial services Deplatforming and application/platform Deplatforming. To do so we ask three questions of the respondents with each question corresponding to a dimension of Deplatforming. Those questions present a hypothetical neutral setting in which the respondents are asked to what extent does a private

company have the right to deplatform a group of interest. The group of interest is either the most liked or least liked group that the respondents ranked in the previous question. Whether a respondent is assigned the most liked group or the least liked group that they picked was randomly decided. Accordingly, the sample was divided into two groups. A group that received three Deplatforming questions with the group that they ranked as their most liked as the group of interest while a second group received the same questions but with the group they ranked as their least liked group of interest. This helps us explore whether people approve or disapprove of Deplatforming in principle or are people's attitudes on Deplatforming dependent on the group in question and the respondents' feelings towards this group.

First, we calculate the average treatment effect for each of the three aspects of Deplatforming and test if the difference in responses between the two experimental groups is significant or not. After we do so, we calculate an additive Deplatforming index that incorporates the three dimensions, and we estimate both the average treatment effect and calculate the significance for that index. It is important to indicate that for this analysis and throughout this paper, the higher the scale, the more acceptance a respondent shows for Deplatforming.

Dimension of Deplatforming	Average Treatment Effect (Mean of Least – Mean of Most)	T-Test p-value	Statistical Significance at the 95% confidence level
Social Media	3.6 - 3.11 = 0.49	Approx. Zero	Significant
Online Financial Services	3.67- 3.06 = 0.61	Approx. Zero	Significant
Application Development	3.56 - 3.18 = 0.38	0.000615	Significant
Deplatforming Index	3.61 - 3.12 = 0.49	Approx. Zero	Significant

Table 2. ATE, p-value and statistical significance indication for the three dimensions of Deplatforming as well as the additive Deplatforming index. The higher the value, the more support for Deplatforming.

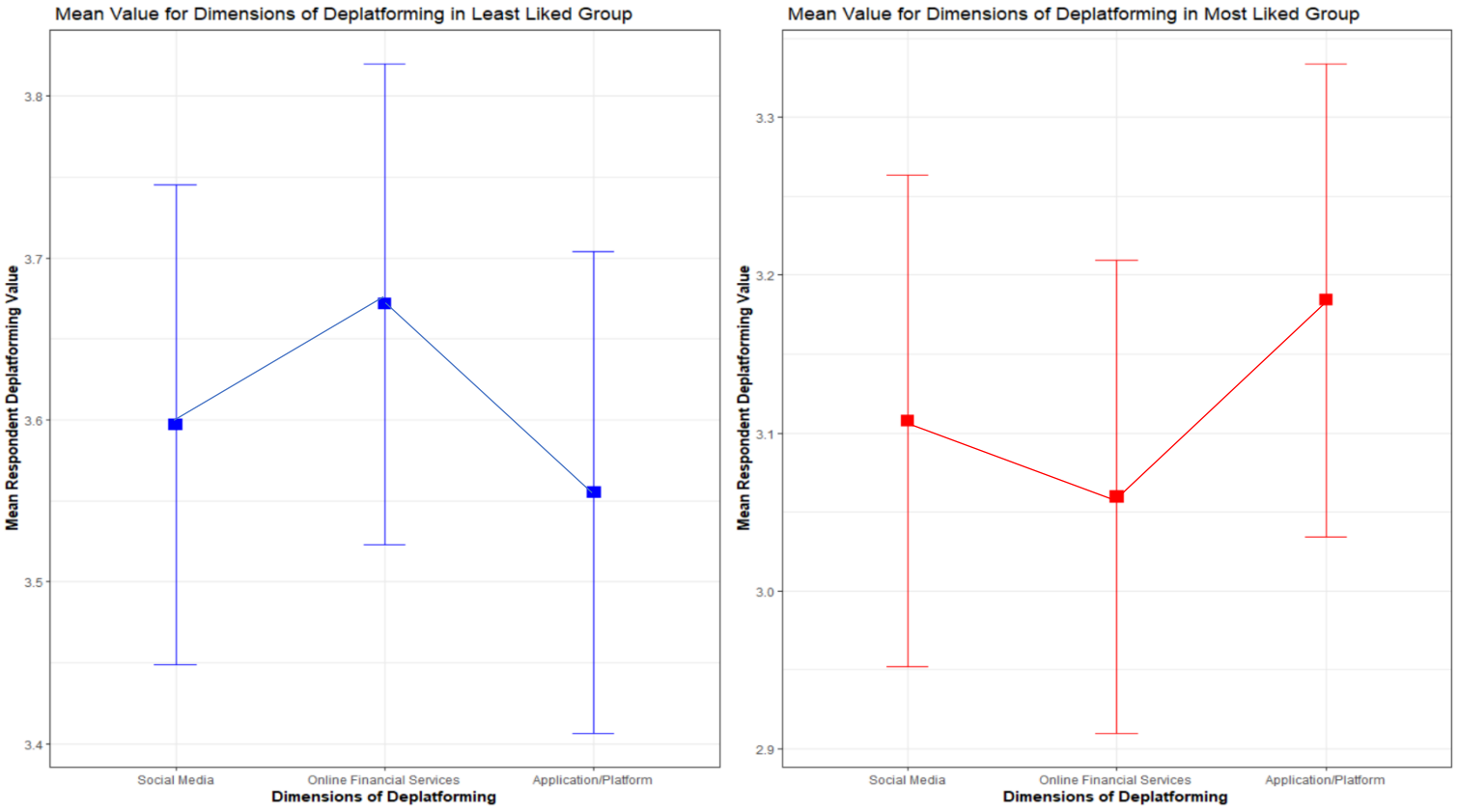
Table 2 shows that the difference between the two experimental groups among all three dimensions of Deplatforming is significant. This indicates that the people become more supportive of Deplatforming on social media when asked about their least liked groups than when asked about their most liked groups with an effect of 0.49. The 0.49 is approximately half a point on the Likert scale which means that when a person is asked about Deplatforming the least liked group instead of the most liked group on social media, they on average become a half a point more accepting of Deplatforming. This effect becomes higher for online financial services Deplatforming (0.61) but lower for Application/Platform hosting Deplatforming (0.38). On average, the effect for the three dimensions is 0.49 which means that on average for the three aspects, people tend to be half a point more supportive of Deplatforming for their least liked groups than their most liked ones.

We have three aspects of Deplatforming with three different effects. In fact, it is interesting that people are more accepting of Deplatforming for financial services while less accepting for

social media Deplatforming and much less for application hosting Deplatforming. One speculation for this discrepancy may be that access to financial services may provide a group with the timely, direct ability to do something in the real world compared to social media and application development which may take time to present activity in the real world. The low score effect for application Deplatforming may also be due to people's reaction to the increasing monopoly of free speech by big tech companies [The Wall Street Journal, 2021]. We can analyze each dimension individually and observe the differences for each dimension between the least liked and most liked experimental groups. However, this report will only focus on the Deplatforming Index.

This brings about a pivotal question. Why do we use an index instead of using the three Deplatforming measures? As observed in Figure 3 below, we can see that the confidence intervals for the three Deplatforming dimensions overlap for both the least liked and most liked experimental groups. This means that those three groups are statistically equal. We also conduct an ANOVA test (not shown) for the three dimensions for both the least liked and most liked groups and we find them significant as well. Since there is no underlying model or assumption for measuring Deplatforming, using an additive index with equal weights is the safest and most popular option. To ease interpretation, the Deplatforming Index was rescaled such that it ranges from 0 (total disapproval/ of Deplatforming) to 1 (total approval of Deplatforming), with 0.5 indicating neutrality in regard to Deplatforming. From Figure 3, we can also observe that there is an inverse trend when comparing the least liked and most liked such that for respondents assigned to the least liked group, as mentioned previously, online financial

services have the highest mean Deplatforming Index value, followed by social media Deplatforming and then application/platform Deplatforming.



(a) (b)

Figure 3. (a) The mean respondent deplatform value for each of the three aspects of Deplatforming but only for the least liked Group. (b) The mean respondent deplatform value for each of the three aspects of Deplatforming but only for the most liked group. Blue is for least-liked and red is for most-liked.

However, for the respondents assigned to the most liked group, we observe the opposite trend with application/platform having the highest mean Deplatforming Index value followed by social media and then online financial services. This tells us that because people are not threatened/concerned enough by their most liked group, they do not fear the consequences of allowing them financial services and thus disapprove more of Deplatforming, unlike the case of

the least liked group. We will now focus our attention to the normalized (0-1) additive Deplatforming index.

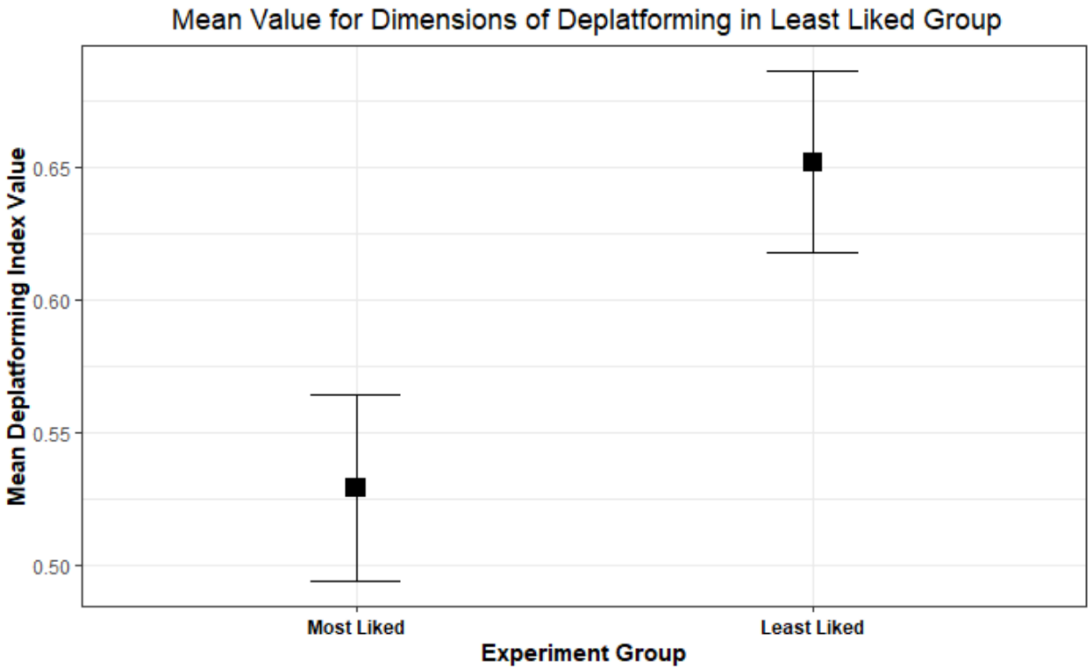


Figure 4. The mean Deplatforming Index value for both the most liked and least liked groups.

Figure 4 above clearly shows the effect that we previously calculate but on a new (0-1) scale reflecting all three dimensions. We observe that on average those in the most liked group have a Deplatforming Index of about 0.53 which is a neutral attitude towards Deplatforming; however, the average Deplatforming Index for the least liked groups is about 0.65 which corresponds to accepting Deplatforming. The effect of the least liked group compared to the most liked group is 0.12 which as just shown is enough to change the average attitude from neutral to more approving of Deplatforming. We also see that the difference is not significant but that was shown earlier in Table 2.

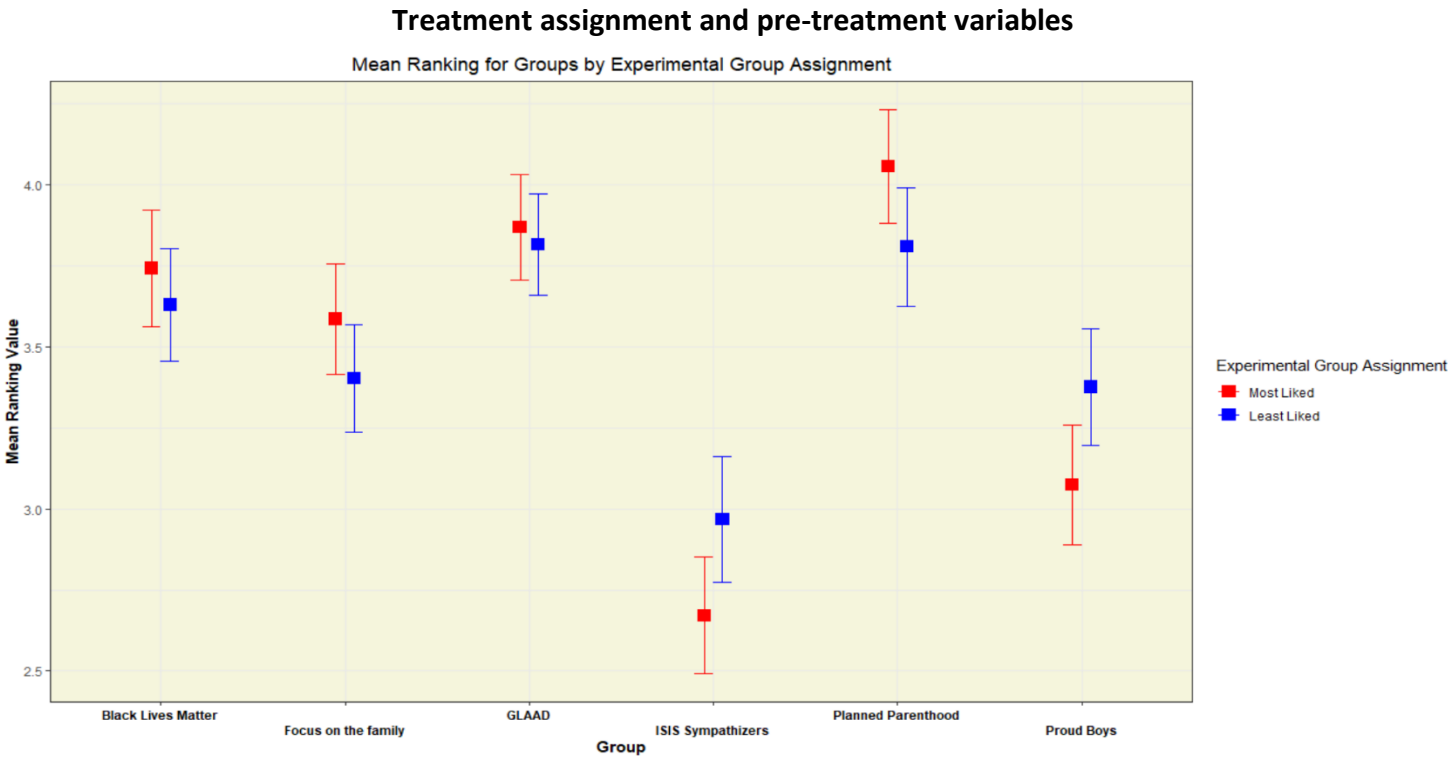


Figure 5. Mean ranking of pre-treatment groups (GLAAD, Proud Boys, etc.) colored by Experiment Group Assignment. Blue is for the least liked group which red is for the most liked group.

As we explored previously, we asked the respondents to rank groups from least to most liked. It is relevant and essential to our analysis to understand what groups had the higher and lower rankings about the two least and most liked experimental groups. From Figure 5, we observe that in general there is no significant difference in ranking between the two experimental treatment groups among all six groups of interest. This indicates that if there is a difference in the Deplatforming Index between experimental treatment groups then that difference is probably attributed to the experiment itself and not pre-treatment group (GLAAD, Proud Boys, etc.) selection. It is also important to note that the pre-treatment groups with the highest and lowest rankings from the pre-treatment analysis we did earlier still holds for both experimental treatment groups. The interesting noticeable trend is those who were assigned to the most liked group treatment rank the top four groups of interest higher than those in the

least liked group treatment: however, this trend reverses for the least two ranked groups, ISIS Sympathizers and Proud Boys. Those who got the least liked assignment ranked those two groups higher than those who were assigned to the most liked treatment group.

The Deplatforming Index and Covariates

In this section we attempt to reexplore (Table 1) the demographic variables and covariates common in literature. We graph the Deplatforming index by those covariates between the group treatment assignment (least/most liked groups) to understand the trend and the significance of the covariate's effects between treatment groups on the Deplatforming Index.

1) Gender

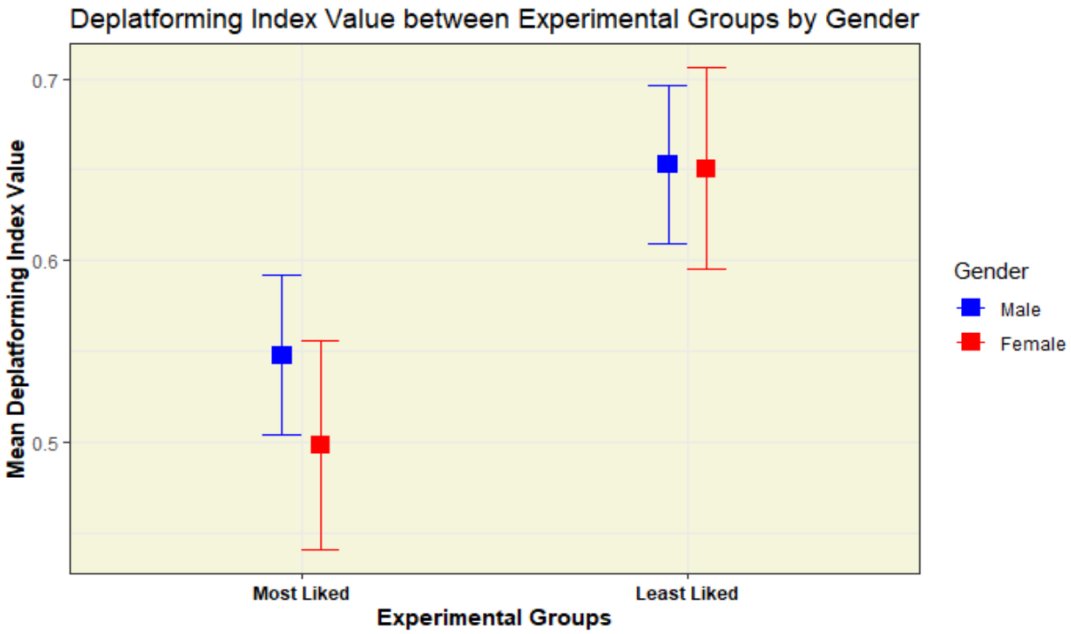


Figure 6. The mean Deplatforming Index value for both the most liked and least liked groups by gender. Points represent mean Deplatforming Index value and color determines treatment assignment. Blue is for male and red for female. We also calculate the confidence interval for the Deplatforming Index (DI)

From Figure 6, we observe that there is no significant difference in the Deplatforming Index between males and females for both least and most liked groups. It is noted that the confidence intervals for females are wider due to their smaller percentage as we discussed

earlier. Also, we notice that the males in the most liked group approve more of Deplatforming (0.55) than females (0.5) while those in the least liked group appear to have approximately the same Deplatforming Index. This means that for the most part there is no significant difference between men and women when it comes to accepting the Deplatforming of most and least liked groups. This contradicts the expectation from literature that Women are usually less willing to tolerate controversial groups than men [Golebiowska, 1999].

2) Age

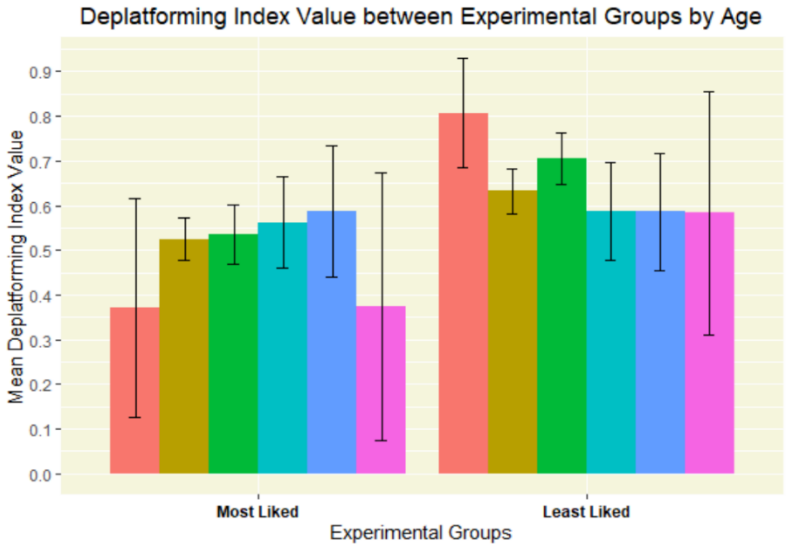


Figure 7. Bars represent the mean Deplatforming Index value for both the most liked and least liked groups by age groups (represented by colors). We also calculate the confidence interval for the Deplatforming Index.

For the Age group graphical analysis, we had to remove the Under 18 category and the 75 to 84 categories since they each had only one respondent and thus incomparable across experimental treatment groups. It also should be noted that there is an 85 and over category but includes zero respondents. Figure 7 shows that when we move from the most liked to the least liked group, we observe an increase in the acceptance for Deplatforming for all age groups, except 35-44 and 45-54, which are approximately equal across and for both groups. The interesting trend is that the increases for other age groups are modest except for the youngest and the oldest age groups. In the most liked group, the oldest (65-74) and youngest (18-24)

respondents had the least accepting for Deplatforming (0.38) which indicates disapproval of Deplatforming and is lower than the most liked group average (0.53). When we move to the least-liked group, however, we observe that youngest age group has a huge Deplatforming index value increase to 0.8, indicating strong acceptance for Deplatforming and higher than the average of 0.65 for the least liked groups. The oldest group also had their index increase but only to 0.6 which is below the average of 0.65. While the difference in effect among most age groups for both experimental groups seem to be insignificant, the difference in effect for the youngest age group between the two experimental groups seems significant. This might make sense because younger people are generally more involved online and thus may hold stronger opinions regarding Deplatforming that other groups. Similarly, we also observe a potential difference between the youngest and the second youngest (25-34) in the least liked group. Also, it is important to note that the differences in confidence levels are due to differences in sample sizes and thus this may create false relationships. These results in part contradict the expectation that the older a person gets, the less tolerant they become [Avery, 1989].

3) Education

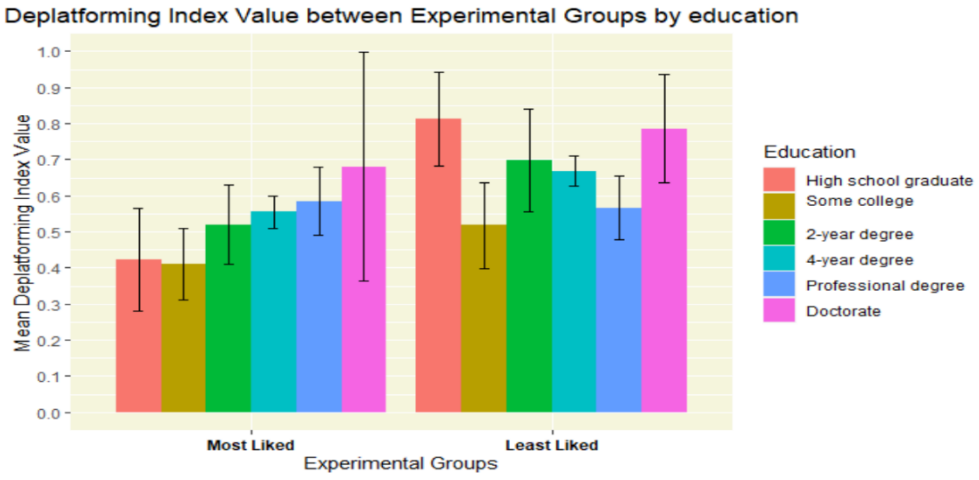


Figure 8. Bars represent the mean Deplatforming Index value for both the most liked and least liked groups by highest completed degree (Education) represented by colors. We also calculate the confidence interval for the Deplatforming Index.

Unlike Age, there is a clear pattern Deplatforming trend for those assigned to the Most Liked group. We can clearly see that as education increases in the most liked group, so does the Deplatforming index, indicating an increasing acceptance for Deplatforming. In the most liked group, those who are high school graduates or went to some college had weak disapproval for Deplatforming (0.4) but as we increased in education, we see that those who hold a Doctorate are more accepting of Deplatforming. The differences among education groups are all insignificant in the most liked group. When we move to the least liked group, we observe that there are potential significant differences within that group, particularly between high school graduates on the one hand and those who went to some college and those with professional degrees, respectively. We also notice that there is no specific trend within the least liked group and that almost all education groups have a higher acceptance of Deplatforming than their most-liked group. Finally, we notice a great increase in the Deplatforming Index when we move from most liked to least liked for high school graduates and those with a doctorate, which is a surprising result since the least and most educated categories have the most acceptance of Deplatforming (approx. 0.8), invalidating any potential theory or trend involving a monotonic/linear relationship between education and Deplatforming when given the least liked treatment. This completely violates the expectation that educated people are more tolerant [Golebiowska, 1995].

4) Conservative/Liberal Leaning

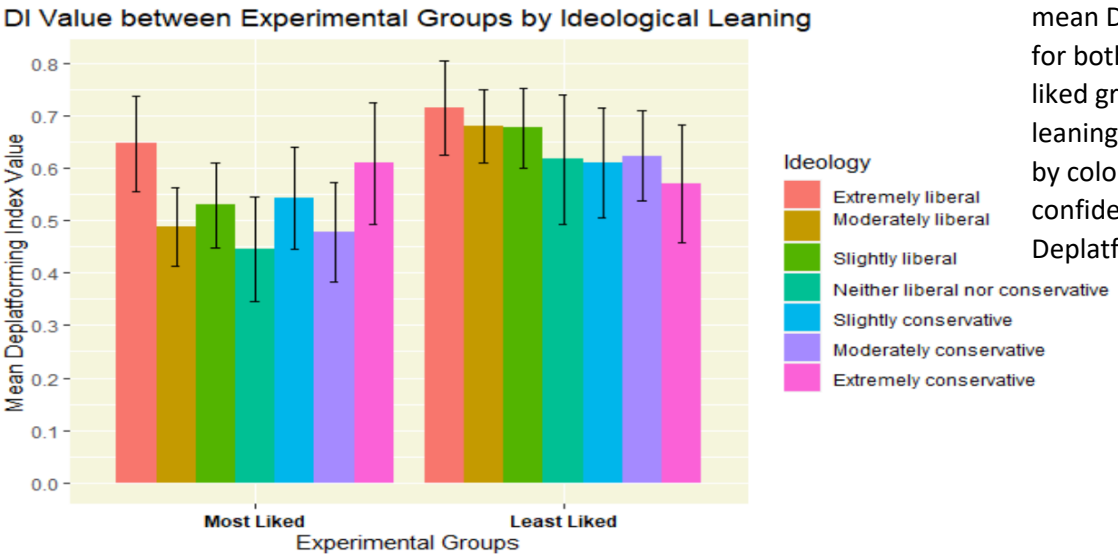


Figure 9. Bars represent the mean Deplatforming Index value for both the most liked and least liked groups by ideological leaning (Ideology) represented by colors. We also calculate the confidence interval for the Deplatforming Index.

Figure 9 shows us that there is no discernible trend or pattern regarding Deplatforming in the most liked group. Those extremely liberal in the most liked group tend to have the highest Deplatforming Index value, indicating acceptance of Deplatforming. The rest of the ideological groups in the most liked group are around or slightly below 0.5 (neutrality) in an alternating fashion. Surprisingly the extremely conservative have a Deplatforming Index value of 0.6 meaning they accept Deplatforming but weakly. It is surprising that both the extremely conservative and the extremely liberal accept Deplatforming to certain degrees while the rest of the groups are on average neutral to Deplatforming. The least liked group tells a different story. First, we observe that all of the ideology categories see an increase in Deplatforming Index when moving to the least liked group. The only exception is the extremely conservative. In fact, the extremely conservative is the only group that becomes *less* accepting of Deplatforming when asked about a least liked group. This contradicts the general trend. It, however, does not contradict the trend *within* the least liked treatment group, since we

observe that the extremely liberal is the most accepting of Deplatforming (0.72) and then the acceptance in Deplatforming decreases as respondents are more conservative till reaching the extreme conservative category which is very weakly accepting Deplatforming and bordering on neutrality (0.55). This is a very interesting finding because it matches the expectation that liberals are more accepting of Deplatforming than conservatives (at least ever since Trump became President) [Pew Research, 2020; The Guardian, 2021]. The only issue, however, that the differences in groups are insignificant. Conversely, there could be significant relationships between categories in the most liked group, namely the extremely liberal and the neither conservative nor liberal.

5) Political Party Leaning

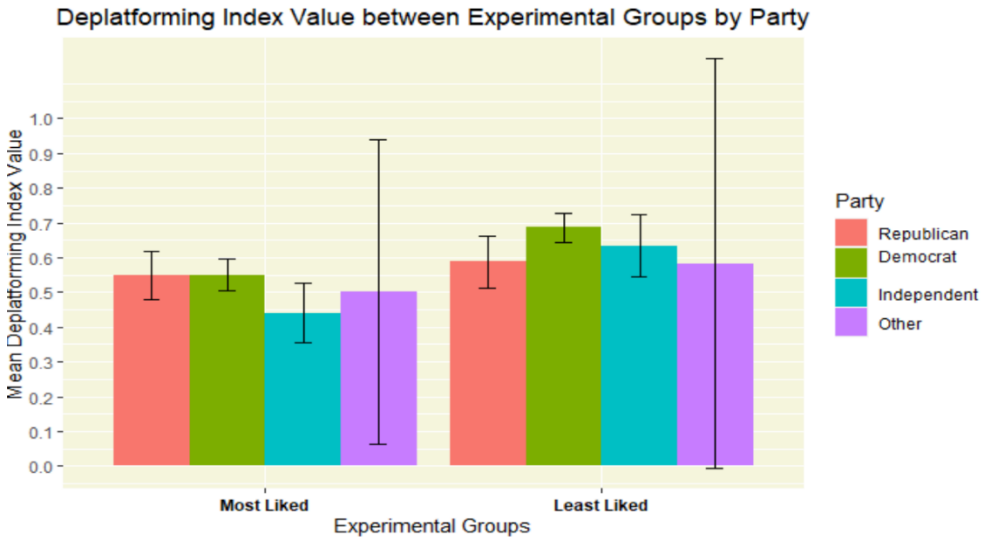


Figure 10. Bars represent the mean Deplatforming Index value for both the most liked and least liked groups by Political Party (Party) represented by colors. We also calculate the confidence interval for the Deplatforming Index.

It is important to acknowledge that the very long confidence interval for the "Other" category for both experimental groups is due to only having 6 respondents in total. Accordingly, the results for "Other" are probably unreliable. The most liked group shows disapproval of Deplatforming (0.4) while Republicans and Democrats are weakly accepting and bordering on neutrality (0.55). When we switch to the least liked group, we observe that the Index value for

Republicans is essentially the same while both Democrats (0.7) and Independents (0.62) see an increase in the Deplatforming Index, resulting in an acceptance of Deplatforming. Generally, for both experimental groups, the categories are statistically equal (insignificant). It has been a trend that Republicans and Conservatives are less accepting of Deplatforming than Democrats and Liberals [Pew Research, 2020; Associated Press, 2021]; however, the effect, graphically at least is insignificant or close to it. We cannot know for sure unless we have a statistical test or utilize regression.

6) Political Activity

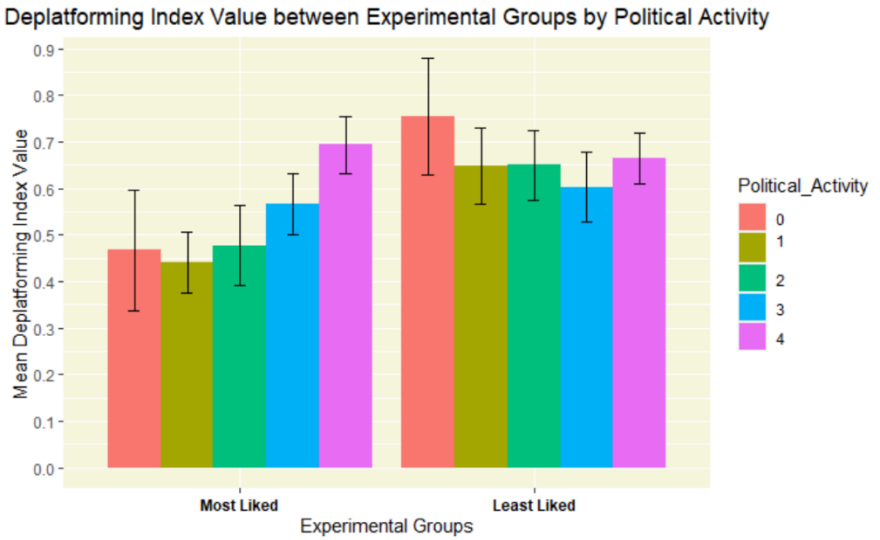


Figure 11. Bars represent the mean Deplatforming Index value for both the most liked and least liked groups by Political Activity represented by different colors. We also calculate the confidence interval for the Deplatforming Index.

Political Activity in our survey is measured by asking the respondents 4 yes/no questions about 4 activities, such as contributing to campaigns or attending rallies, to reflect their involvement in politics in the past year. For the purposes of this analysis, we decided to categorize those questions into one variable such that people who did not engage in any kind of political activity get a score of 0, those who engaged in one activity get a score of 1, etc. Figure 11 provides an interesting result such that not only do we observe a pattern across both experimental groups but we also notice the opposite patterns. For the most liked group, we observe that as people

get more politically active, they become more accepting of Deplatforming. Such that those not involved into any political activity are approximately neutral (0.5) to Deplatforming which makes sense but as a person gets politically involved may be they better understand the rights and power of private companies and corporations and thus become more accepting of Deplatforming groups, especially if those groups (most liked) are non threatening. However, the alternative happens when the group (least liked) involved is threatening since those political involved people may want to extend protection especially to those who are deemed controversial because it happens a lot nowadays. Regardless of all of that, it should be noted that we are analyzing the variation within each experimental group and not comparing those groups. If we do so, we immediately realize that most of the categories least liked experiment group have a higher Deplatforming Index compared to their most liked group counterparts. Another observation would be that the increase in Deplatforming index from most to least liked experimental groups is most obvious across those who are less politically active (0 to 2). Those groups have an increase of 0.15 to 0.25 points on the Deplatforming index scale which easily moves them from neutral/disapproving of Deplatforming to accepting of Deplatforming. This is contrasted with people who are more active (3-4) who only see a 0.05-point increase in the Deplatforming index at the most. While the categories least liked groups are statistically equal (insignificant). Those who are engaged the most in political activity (4) may be significantly different from the rest of the categories, however this is the case only in the most-liked group.

We selected Political Activity as an alternative to the widely used “Political Knowledge” variable used in literature as a determinant of tolerance. This report attempts to observe if knowledge and actual activity is interchangeable or if the trend is actually different. The

literature expectation is that more Political Knowledge increases tolerance among people [Hall, 2017]. Interesting enough, the trend does not hold for the safer, most liked groups but weakly holds for the "more dangerous" least liked group.

7) Ethnicity

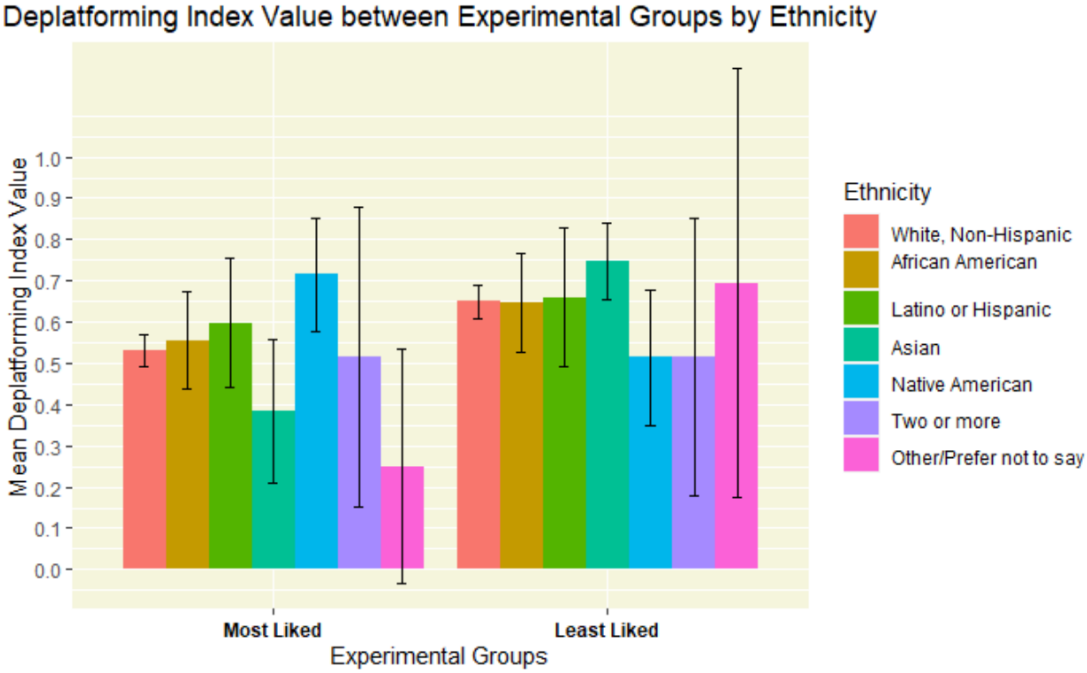


Figure 12. Bars represent the mean Deplatforming Index value for both the most liked and least liked groups by Ethnicity represented by different colors. We also calculate the confidence interval for the Deplatforming Index.

Figure 12 shows the differences in attitudes towards Deplatforming according to different Ethnicities. In the most liked group, we observe that Native Americans have the highest Deplatforming index value (0.7), indicating an acceptance of Deplatforming. Asians, on the other hand, have the lowest value and this show a lack of acceptance/support towards Deplaforming (0.4). The trend, however, is surprisingly reversed when observing the least liked group. When asked about a less liked group, Asians now have the highest Deplatforming index value (0.75) while Native Americans are more neutral (0.5). This indicates that Asians do not accept Deplatforming when its group they prefer but strongly do so when it is a group they may dislike. The other three groups, White, African American and Latino, do not show any particular

trends. Only that their Deplatforming acceptance increases when we move from the most liked to the least liked group. We do not review or analyze the Two or More and the Other category since they contain very few responses, as evidenced by the large confidence intervals. Additionally, there does not seem to be a significant difference between groups; however, Asians may be significantly different not only across both groups but even within each group in relation to other ethnicities.

8) Religion

The effect of religion on political tolerance is usually not extended to non-Christian religions due to their minority status in the US. It is usually concentrated on Christian denominations [Katnik, 2002; Schwadel et. al, 2018]. We try to do something novel and analyze the difference across different religions. The problem is that looking back at the demographics in Table 1, we find that the sample is predominately Christian and the remaining of the sample is mostly nonreligious. This means this question will only be an indirect proxy to the "degree of religiousness" question that we already have for this analysis. Consequently, we skip the graphical analysis for this question and acknowledge the design flaw in this question. This validates other analyses that focus on Christian denominations since they provide more substantive results.

9) Degree of Religiousness

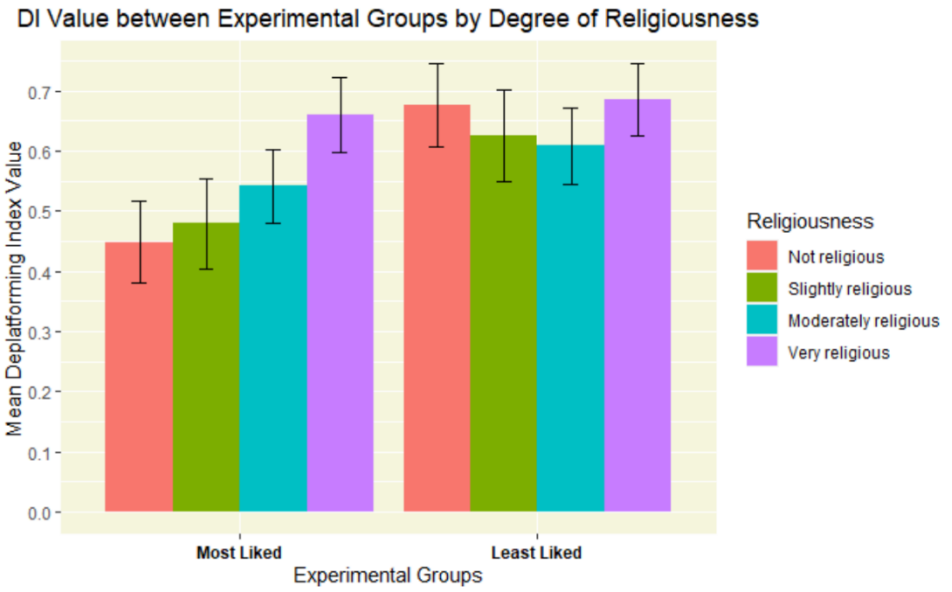


Figure 13. Bars represent the mean Deplatforming Index value for both the most liked and least liked groups by Degree of Religiousness (Religiousness) represented by different colors. We also calculate the confidence interval for the DI.

In political tolerance literature, it is observed that the more religious a person is, the more intolerant they tend to be [Eisenstein, 2006]. The respondents in the Most Liked group show exactly that trend such that the more religious a respondent is, the higher Deplatforming value we observe. We move from a (0.45) Deplatforming value indicating weak disapproval of Deplatforming for nonreligious respondents to an acceptance of Deplatforming value of 0.65 for the very religious. Interestingly enough, when we move to the least liked group, we observe that nonreligious people tend to have an approximate 0.2 increase in the Deplatforming index which moves them from weakly disapproving (0.45) to moderately accepting Deplatforming (0.65). This change shows that the nonreligious have an attitude towards Deplatforming similar to that of the very religious if both were asked about their least liked group. This observation contradicts the contrast between the religious and nonreligious observed in literature. We also observe that the very religious stick to their Deplatforming (value) stance regardless of the experimental group. This implies that they may support Deplatforming in principle which is

supported in literature [Eisenstein, 2006]. Finally, we see that for the most part the differences between groups are insignificant.

Regression Analysis

As previously observed, sometimes it is difficult to conclude statistical significance solely from plots. However, the more important issue is that by using graphs we only explore the individual effect of the covariate of interest on the Deplatforming Index given the treatment assignment. Controlling for other variables is a sure way to produce more certain statistical effects and correlations, holding all other variables constant or equal. Since our outcome of interest (Deplatforming Index) is a continuous measure, using linear regression is one of the more popular and straightforward methods to establish which controls exactly affect the respondents' Deplatforming Index scores. Since we regress the Deplatforming Index on the treatment assignment, all the covariates of interest explored in this paper (most of which are categorical) and their interaction, we get over 50 regression terms. Accordingly, we will only provide the coefficients and interpretation for the terms that are at least 95% significant. The regression equation is as follows.

$$y_i = D_i + \beta_0 + \sum X_{ij}\beta_j + \sum X_{ij}\beta_j * D_i + \epsilon_i, \quad \text{where}$$

y_i is the continuous Deplatforming index value for a respondent i

D_i is the binary (0/1) treatment assignment value for a respondent i such that 1 indicates that the respondent is assigned to the least liked group and 0 indicates the most liked group.

X_{ij} represents the value for a covariate j for a respondent i. Note that a covariate j may include several dummy variables if it is a categorical covariate.

Regression Table

Variable Name	Coefficient	Standard Error	P-value	Significance
Gender (Male)	0.091022	0.038175	0.0174	95%
Moderately Conservative	-0.143167	0.072836	0.0489	95%
Political Activity	0.035798	0.017479	0.0410	95%

Table 3. A summary of the names of the significant covariates, their coefficients, standard errors, p-values and statistical significance level indication.

We only observe 3 significant variables. None of those 3 variables is an interaction between treatment assignment and a covariate. It needs to be highlighted that we know (and analyzed just to make sure) that the treatment assignment, without adding it as an interaction, is in fact significant due to the difference in effect we observed earlier. Not having a significant interaction even though the treatment assignment is significant indicates that the effect due to the treatment assignment cannot be explained/predicted by the determinants we currently have or that we may need a larger and more representative sample to capture the differences in small-sized, under sampled categories. This addresses our question of whether determinants of political tolerance are appropriate to measure the concept of Deplatforming. The interpretation of the significant variables is as follows.

First, we observe that Males have a Deplatforming Index with 0.09 more points than Females, controlling for other variables. This is about one tenth of the index and contradicts the political intolerance literature expectation that men are more tolerant than women especially

towards controversial groups [Golebiowska, 1998]. Second, we observe that people who are moderately conservative have a Deplatforming Index with -0.14 less points than the extremely liberal, controlling for other variables. This conforms to the current expectation that conservatives oppose internet censorship and Deplatforming [Pew Research, 2020]. Third, as the measure of political activity increases by one (a person engages in an additional political activity), the Deplatforming index increases by 0.036, controlling for other variables. This means that as a person becomes more engaged in political activity they tend to accept and approve more of Deplatforming.

Conclusion

We had three research questions that we attempted to tackle using a survey experiment. First, we showed that there is indeed a significant effect (0.12) between the two experiment groups such that respondents were prone to be more accepting of Deplatforming when they were asked about a group, they liked least than those who were asked about a group they liked most. We then observed that there was no significant difference between the three dimensions of Deplatforming and thus decided to use an equally weighted additive index to represent Deplatforming. Using the theoretical determinants of political tolerance, we attempted to individually explore how those determinants affect the index through treatment assignment. We noticed trends that were contradictory to literature such as younger people being more accepting of Deplatforming than older people when asked about least liked groups, more educated groups tend to be more accepting of Deplatforming when asked about most liked groups and even least liked groups (this only holds for Doctorate holders) and the most and least religious being equally accepting of Deplatforming for the least liked groups. We also

observed trends that were on par with political tolerance literature such as liberals and democrats being more accepting of Deplatforming than conservatives and republicans, respectively. We also observe that Asians are much more accepting of Deplatforming for least-liked groups but less accepting for most-liked groups; however, Native Americans exhibit the opposite trend. Finally, we do a regression analysis to account all the determinants and statistically determine which variables explain the variation in the Deplatforming index both independently and through the treatment assignment. We did not observe any significant interaction between the treatment assignment and the covariates but only 3 main effect variables that include the gender of the respondents, whether or not the respondents are moderately conservative and political activity of the respondents. Being male and being more engaged in political activity both, respectively, correlates with more acceptance of Deplatforming while being moderately conservative correlates with less acceptance. The report concludes with two thoughts. A large, representative sample is always the best way to be able to provide robust and indicative results. Finally, it seems that Deplatforming may need to be analyzed from a different lens than political tolerance or at least more research into the determinants of Deplatforming need to be investigated. This is evidenced by the fact that one of the few variables that were significant, Political Activity, was added to the report as an alternative to the political expertise variable traditionally used in political tolerance literature.

References

1) Bokat-lindell, S. (2021, January 14). Deplatforming Trump could WORK. but at what cost? Retrieved April 10, 2021, from https://www.nytimes.com/2021/01/14/opinion/trump-deplatform-twitter.html
2) Twitter 'permanently suspends' Trump's account. (2021, January 09). Retrieved April 11, 2021, from https://www.bbc.com/news/world-us-canada-55597840
3) Merriam-Webster. The Words of the Week - 8/10/2018. Merriam-Webster. https://www.merriam-webster.com/words-at-play/the-good-the-bad-the-semantically-imprecise-08102018.
4) Rogers, R. (2020). Deplatforming: Following extreme Internet celebrities to Telegram and alternative social media. European Journal of Communication, 35(3), 213–229.
5) Duffield, W. (2021, January 25). A brief history of 'deep deplatforming'. Retrieved March 11, 2021, from https://www.cato.org/blog/brief-history-deep-deplatforming
6) Sullivan, J., Piereson, J., & Marcus, G. (1979). An Alternative Conceptualization of Political Tolerance: Illusory Increases 1950s–1970s. American Political Science Review, 73(3), 781-794.
7) Boch, A. (2020). Increasing American Political Tolerance: A Framework Excluding Hate Speech. Socius.
8) Clawson R. A. & Oxley Z. M. (2008). Public Opinion: Democratic Ideals, Democratic Practice. International Journal of Public Opinion Research.
9) Gibson, J., Claassen, C. & Barceló, J. (2020). Putting Groups Back into the Study of Political Intolerance.

10) Stouffer, S. (1955). Communism, Conformity and Civil Liberties. New York: Doubleday.

11) Needleman, S., & Wells, G. (2021, January 11). Twitter, facebook and others silenced trump. now they learn what's next. Retrieved April 13, 2021, from https://www.wsj.com/articles/twitter-facebook-and-others-silenced-trump-now-they-learn-whats-next-11610320064?mod=rss_TechnologyGolebiowska, E. (1999). Gender Gap in Political Tolerance. Political Behavior, 21(1), 43-66. Retrieved April 10, 2021, from http://www.jstor.org/stable/586585

12) Avery, P. (1989). Adolescent Political Tolerance: Findings from the Research and Implications for Educators. The High School Journal, 72(4), 168-174. Retrieved April 13, 2021, from http://www.jstor.org/stable/40387546

13) Golebiowska, E. (1995). Individual Value Priorities, Education, and Political Tolerance. Political Behavior, 17(1), 23-48. Retrieved April 13, 2021, from http://www.jstor.org/stable/586556

14) Vogels, E., Perrin, A., & Anderson, M. (2020, September 18). Most Americans think social media sites censor political viewpoints. Retrieved April 9, 2021, from https://www.pewresearch.org/internet/2020/08/19/most-americans-think-social-media-sites-censor-political-viewpoints/

15) Claim of anti-conservative bias by social media firms is baseless, report finds. (2021, February 01). Retrieved April 13, 2021, from https://www.theguardian.com/media/2021/feb/01/facebook-youtube-twitter-anti-conservative-claims-baseless-report-finds?CMP=oth_b-aplnews_d-1

16) Izaguirre, A. (2021, March 07). GOP pushes bills to allow social media 'censorship' lawsuits. Retrieved April 13, 2021, from https://apnews.com/article/donald-trump-legislature-media-lawsuits-social-media-848c0189ff498377fbfde3f6f5678397

17) Hall, J. P. (2017). Effects of political knowledge on political tolerance. Journal of Political Science Education, 14(1), 104-122. doi:10.1080/15512169.2017.1366326

18) Katnik, A. (2002). Religion, social class, and Political Tolerance. International Journal of Sociology, 32(1), 14-38. doi:10.1080/15579336.2002.11770242

19) Schwadel, P., & Garneau, C. R. (2018). Sectarian religion and Political tolerance in the United States. Sociology of Religion, 80(2), 168-193. doi:10.1093/socrel/sry032

20) Eisenstein, M. A. (2006). Rethinking the relationship between religion and Political tolerance in the US. Political Behavior, 28(4), 327-348. doi:10.1007/s11109-006-9014-5