The Mechanisms of Protest Recruitment through Social Media Networks*

Andreu Casas Joshua Tucker Leon Yin Jonathan Nagler
New York University

Jennifer M. Larson
Vanderbilt University

DRAFT

Abstract

The literature on protest mobilization has long suggested that social ties have a strong influence in the decision to protest. Recent literature on social media mobilization also shows that in the current digital environment these social ties effects often take place online, particularly on social media. However, previous research does not provide clear evidence for why personal ties play a mobilizing role. In this paper we lay out four main theoretical mechanisms and we test them using real-world protest attendance data: social networks mobilize because a) they provide basic logistic *information* that is vital to protest coordination, and b) they create *motivations* for others to protest, c) they solve *coordination* problems, and d) they put *pressure* on others to participate. We collect data on Twitter activity during the 2018 Womens March that took place in many cities in the United States on January 20th. We then use geolocated accounts to find a set of users who attended a march and a set of users who did not. We use machine learning techniques to determine the amount of information, motivation, coordination, and pressure messages to which they were exposed through their Twitter networks. In line with current theories, we find users who protested to be more connected among themselves than those who did not. In regards to the mechanism analysis, we find that, even when controlling for potential confounders, users who saw at least one information and coordination message were more likely to protest, but find a null effect for motivation and pressure messages.

*The authors thank the members of the Social Media and Political Participation (SMaPP) lab for early feedback on the project, as well as Emily Constance Moffa, Shiva Darshan, Megan Brown, and Nick Kuegel for their assistance. A previous version of the paper was presented at the Annual Meeting of the Comparative Agendas Project, July 4-6, Amsterdam. We would also like to thank and knowledge the financial support of the Moore and Sloan Foundations, the Bill and Melinda Gates Foundation, and the John S. and James L. Knight Foundation. Corresponding author: Andreu Casas (andreucasas@nyu.edu).

1 Introduction

On January 21st 2017, more than four million people joined the Women's March that took place in hundreds of cities around the United States to demand a more equal and safe society for women.[1] In conjunction with other movements and events that happened right after (such as the #MeToo and the #TimesUp movements), the march contributed to building a stronger coalition around gender equality and to increasing the salience of existing gender disparities – illustrating the ability of social mobilizations to set political agendas and eventually influence policy. How did this protest become to be the "largest demonstration in US history" (Frostenson, 2017)? But more importantly, what influences people's decision to protest?

Social scientists have long debated this question. A long-standing argument is that the decision to protest is in large part a function of the activity of others in one's social network. Put another way, people are more likely to attend a protest if people with whom they have social ties attend as well – a claim supported by findings based on self-reported data (Opp and Gern, 1993), real-world behavior (Larson et al., 2018) and both (McAdam and Paulsen, 1993).

However, why do social ties play such a mobilizing role, and through what channels? Scholars have argued that personal networks influence one's decision to protest because they: a) provide basic logistic *information* needed for attendance (Olson, 1965), b) present reasons and a *motivation* for others to attend (Katz and Lazarsfeld, 1955a), c) contribute to solving *coordination* problems (Kuran, 1995), and d) put *pressure* on others to participate (Gerber et al., 2008). Nevertheless, empirical research using real-world data to explore the validity of these mechanisms is nonexistent. Moreover, recent research points to social

[1] There were demonstrations in other countries around the World, but the 4-5 million estimate is only for the United States. The following news article in Vox.com contains information about the head count: *The Women's Marches may have been the largest demonstration in US History*.

media as a channel through which these network effects could be transmitted (Gonzalez-Bailon et al., 2011; Barbera et al., 2015; Jost et al., 2018; Larson et al., 2018; Langer et al., 2018).[2] The extant literature, however, lacks systematic analysis of the *content* of these social media messages, instead largely focusing on measures of networks and network ties, thus leaving questions of *why* and *how* personal ties influence people's decision to protest mostly unanswered.

Here, we test four different hypotheses – one for each of the mobilizing mechanisms described in the preceding paragraph – regarding the relationship between the content of messages to which potential protesters were exposed in the days leading up to the Women's March that took place on January 20th 2018, and participation in the march. More specifically, we use geolocated Twitter accounts to find a set of Twitter users who attended one of the 249 marches in the United States that day: we refer to them as the *Protester* group. As comparison set, we identify Twitter users who could have but did not attend one of the marches. These are users who also geo-located their tweets, also tweeted about the Women's March, and did so from an area that was close enough to a march (and during the time of that march) that they could have plausibly participated. We refer to them as the *Non-Protester* group.[3] This allows us to test whether the protester and non-protester networks vary in predictable ways (Larson et al., 2018). To test the empirical support for the four hypotheses regarding the *mechanisms* by which exposure to tweets may have impacted the protest decision, we also collect all the march-related tweets that both our protesters and non-protesters could have seen the week before the march. After training machine learning classifiers to predict the presence of the four mobilizing mechanisms in each of the messages (information, motivation, coordination, and pressure), we can then assess whether people

[2] A weaker claim, but no less relevant, would be that online social networks serve as a good proxy for studying network effects that may occur both online and offline.

[3] To probe the robust of our findings, we also control for a number of other covariates that could distinguish the two groups; see **??** for more detail.

exposed to the four mobilizing message types on Twitter were more likely to protest.

We first find that, although, users in the non-protester set had on average a larger number of social media ties (friends), ties of ties (friends of friends), reciprocated ties (friends who follow the user back), and triadic ties (friends among their friends' friends), protesters were indeed more connected than non-protesters among each other: they had a larger proportion of ties, ties of ties, reciprocated ties and triadic ties within their own (protester) group.

Then, when exploring the messages that protesters and non-protesters's networks exchanged the week before the march, we find that exposure to an *information* and also to a *coordination* message increased people's propensity to protest. We find a null effect for *motivation* and *pressure* type messages. The findings are robust to the effect of other confounders.

The contribution of the paper is two-fold. First, building on Larson et al. (2018)'s strategy, we use real-world network and participation data to empirically test the validity of the network-mobilizing effect (Granovetter, 1978; Marwell et al., 1988; Siegel, 2008; Centola, 2013) in the context of a highly decentralized (and somewhat divisive) mobilization. Importantly, we then bring the analysis one step forward and test hypotheses regarding why social networks can be mobilizing. This is the first study up to date that combines real world protest attendance, information about participants and non-participants' networks, and the messages these networks exchanged the days previous to protest. Overall, the paper provides an exhaustive picture of the conditions under which personal networks influence people's decision to protest.

The rest of the paper proceeds as follows. First, in the next section we present the logic behind the network mobilization argument. Then, we build on a wide range of theories to advance a set of four mechanisms that could explain the mobilizing role of networks. We then discuss the data and methods used in the study, and we present the results. We conclude by summarizing the findings and its implications, and by laying out potential next steps for future research.

2 The Mobilizing Role of Networks

An extensive social science literature finds that personal networks play a crucial mobilizing role in protest participation (Granovetter, 1978; Marwell et al., 1988; Siegel, 2008; Centola, 2013; Larson et al., 2018). Some focus on the contagion effect of networks: information about a given collective action can travel more easily through existing personal connections, making those connected to people who are central to the organization more likely to participate than others (Marwell et al., 1988; Centola and Macy, 2007); while others study the interaction between network characteristics and exposure, and people's propensity to protest: given that not all people are initially equally likely to participate in a given action, what kind of personal interactions can make them reach their individual participation 'thresholds' (Granovetter, 1978; Chwe, 2000)?

As Larson et al. (2018) point out, this existing literature makes two main claims. The first one is that personal ties matter. As an example, Opp and Gern (1993) find that participants in the 1989 Leipzig rebellion were more likely than non participants to recall having a tie with other people that protested. The second main point is that the quality or strength of the personal ties are also important. To give one example, Centola and Macy (2007) find that the "width" of a tie (the number of individuals connecting two people) facilitates complex diffusion, such as the spread of high-risk social movements. As another example, a study of participation in another high-risk action, the 1964 Mississippi Freedom Summer Project (where students went to the South of the United States to combat racial seggregation), McAdam and Paulsen (1993) also find that people who applied for and joined the program had stronger ties (Granovetter, 1978) than other participants who applied but did not show up.

Past research on the mobilizing role of social ties has had to rely on self-reported information from those who attended (and did not attend) a protest action, and/or had limited data on the composition of the networks of the people being studied, significantly limiting the scope of the analysis and the generalizability of the findings. The emergence of networked technologies and social media, however, has transformed the study of networks and collective action (Gonzalez-Bailon et al., 2011; Barbera et al., 2015; Larson et al., 2018). Some social media studies focus on the diffusion of a protest movement online in order to pin down what network characteristics allow for their online diffusion (Gonzalez-Bailon et al., 2011; Barbera et al., 2015). More recent research goes one step forward and uses geolocated social media accounts to combine rich network data with real-word protest participation. As an example of the latter, Larson et al. (2018) study the Twitter networks of a group of users who attended the 2015 Charlie Hebdo protest in Paris, and a group of users who were supportive of the action but decided not to attend, and find the group of protesters to be more connected among themselves than the non-protester group. Protesters had a higher proportion of ties, ties of ties, reciprocated ties, and triadic ties within the group.

We build on Larson et al. (2018)'s method and use Twitter geolocated accounts to detect a set of users who attended, and a group of users who despite showing interest did not attend, the 2018 Women's March in the United States. The novel contribution of the paper is to go one step forward and to explore the mechanisms through which personal networks mobilize protest participation. As an additional contribution towards accumulating scientific knowledge, we replicate the analysis of Larson et al. (2018), a study of protest participation in a single national protest in France, to see if we get similar findings in another country – the United States – and context – multi-city protests – to see if similar results emerge (they do).

2.1 Modeling Network Participation Effects using Twitter

Consider a set of people whose political views and interests align with the purpose of a protest. Building on the described literature, we assume that prior to protest not all individuals value attending the protest equally nor have been exposed to the same amount of logistical and motivational content: a set of them value the protest highly, have detailed logistical information on how to participate, are well aware of the motives behind the protest, and will attend independently of what others do. However, others' valuation is a function of their own private reasons for attending, plus an interest in winning the favor of the friends who care about the protest, plus exposure to the necessary logistical and motivational content needed for participation. This means that a higher exposure to others' intentions, motivations, and information should increase one's valuation and willingness to protest.

Exposure to someone' intention, information, and motives is mainly determined by two factors: distance and strength. By *distance* we mean the number of existing personal connections (ties) that would need to be activated in order for a person A to be connected to another person B and to learn about her intentions, information, and motives. The larger the number of intermediate ties between A and B, the lower the probability that A will be exposed to B. By *strength* we mean the quality and the number of existing personal connections between A and B. The quality of the connection between A and B will be stronger if A knows B but also if B knows A (rather than if only A knows B). The connection will also be stronger if A and B have other connections in common (e.g. C). The stronger the tie between A and B, the higher the probability that A will be exposed to B's intentions, motives and information.

If influence-by-exposure is indeed at play, protesters should then be embedded in networks where people are highly exposed to one another. On the contrary, potential protesters who did not attend should be embedded in networks where exposure to one another is lower. Given a set of protesters and non-protesters, a key prediction of this argument is that people in the protester group will be more connected to one another than people in the non-protester group.

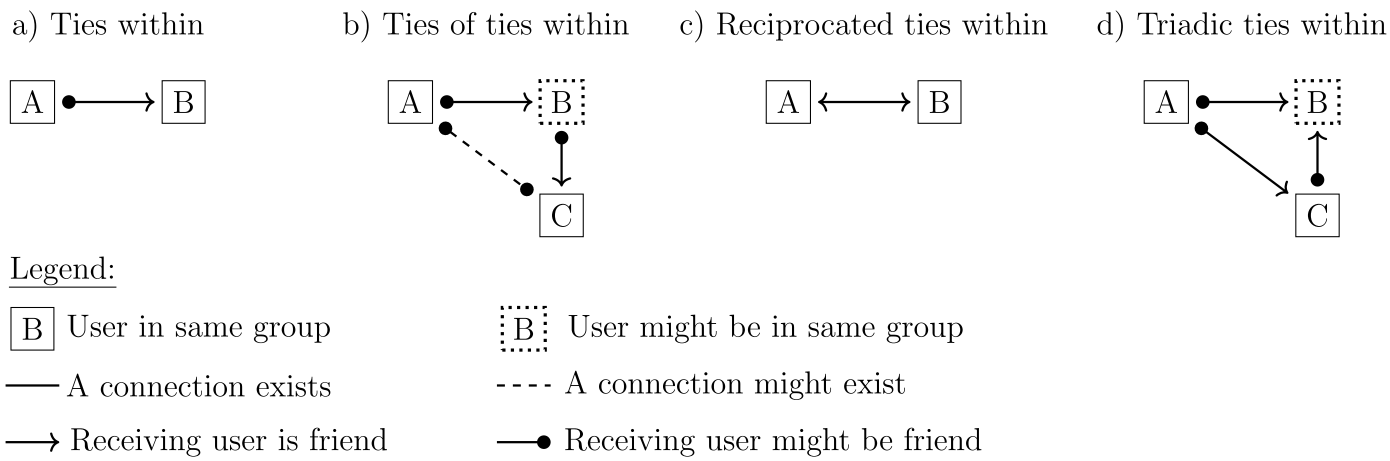
We use Twitter data to test four hypotheses related to this argument. The first two hypotheses (H_1 and H_2) and network measures of interest concern *distance*: the **ties** and the **ties-of-ties** that Twitter users in each (protester and non-protester) group have *within* their own group. We remain agnostic about whether these users were exposed to others' intentions to participate via messages in the social media platform, or via offline face-to-face conversations.[4]

Twitter users can follow other people in the social media platform, and by doing so, they expose themselves to these "friends"'s messages. This following behavior represents the shortest *distance* between two users: at least one of them knows the other one directly and can easily be exposed to her intentions, motives, and information. We consider the users followed by a given user as her ties. These ties, and the rest of the measures, can be easily visualized in Figure 1.[5] As shown in Figure 1a, B is a tie of A because she is followed by A (indicated by an arrow in B's end of the edge). In this case we do not take into consideration whether B follows A back (indicated by a dot in A's end of the edge). We are interested in evaluating whether protesters represented a group where people were more highly exposed to one another (compared to the non-protester group), and apart from measuring the total number of ties the users in each set have, we particularly measure the ties that users have within the same (protester or non-protester) group (indicated in Figure 1a by a solid line

[4]Twitter networks and communications are a fair representation of people's real-world networks and interactions (Bisbee and Larson, 2017), which means that both scenarios are feasible and non-mutually exclusive, and that Twitter data provides a great opportunity to test the network participation argument without making explicit assumptions about the channel through which these network effects took place.

[5]In Figure 1, a solid edge indicates a connection between two users (and a dotted edge indicates that a connection may or may not exist between two users), an arrow at the end of the edge indicates that the user is followed by the user on the other edge (and a dot, instead of an arrow, indicates that the user may or may not be followed by the user on the other edge end), and a solid line around a given user indicates that the user is part of the same (protester or non-protester) group to which the user on the other side of the edge is also part of (and a dotted line around the user indicates that she may or may not be part of the same group).

Figure 1: Visual representation of the network measures we use to test the network mobilization argument.



around each user). Building on the described literature on network mobilization, we expect to find that:

H$_1$: *On average, the proportion of the protesters' **ties** that are also in the protester group is greater than the proportion of the non-protesters' ties that are in the non-protester group.*

The next shortest distance between Twitter users is a situation in which two users (e.g. A and C in Figure 1b) are connected through an intermediate tie (B): only one existing personal connection needs to be activated in order for A to learn about C's intentions, motives, and information (user B needs to share them with A). We take into consideration the ties of ties of users in our protester and non-protester groups, and we particularly care about the ties-of-ties that they have within their own group (notice the solid lines around both A and C in Figure 1b). We apply no specific restriction in regards to whether the intermediate tie B is also part of the (protester or non-protester) group (notice the dotted line around B in Figure 1b), nor whether C is also a tie of A (notice the dotted edge and the dots in both ends of the A-C edge in Figure 1b). In line with what we stated in H_1, we also expect to find that:

H_2: *On average, the proportion of the protesters'* ***ties-of-ties*** *that are also in the protester group is greater than the proportion of the non-protesters' ties-of-ties that are also in the non-protester group.*

The next two hypotheses (H_3 and H_4) and measures of interest concern *strength*: the **reciprocated ties** and **triadic ties** that Twitter users in each (protester and non-protester) group have *within* their own group. First, we consider a user's (e.g. A's) tie within her own group (B) to be stronger if this user is also a tie of her tie (if B follows A back on Twitter: see the arrow in both sides of the edge in Figure 1c). Finally, we also consider a tie to be stronger if a user A shares a tie B with a tie C within her group. As one can see in Figure 1d, this common tie B may or may not be in A and C's (protester or non-protester) group, and C does not need to follow A back and B does not need to follow either of them. In sum, exposure to the intentions of stronger ties to participate should have a more substantive participation effect on people, and we hence expect that:

H_3: *On average, the proportion of the protesters'* ***reciprocated ties*** *that are also in the protester group is greater than the proportion of the non-protesters' reciprocated ties that are also in the non-protester group.*

H_4: *On average, the proportion of the protesters'* ***triadic ties*** *that have one tie in the protester group is greater than the proportion of the non-protesters' triadic ties that have one tie in the non-protester group.*

3 The Mechanisms of Network Mobilization

Hence, an extensive literature finds that people are more likely to join a protest movements if they have ties to other protesters – and even more if these ties are 'stronger.' A crucial question that remains unanswered however is: why is this the case? What is about personal networks that make people more likely to protest? What specific functions do networks play?

Literature from a variety of fields provide several answers this question. Empirical research disentangling the validity of these mechanisms however is nonexistent. This is in part due to data limitations faced by previous studies. Disentangling and/or adjudicating between mechanisms requires to find people who participated in a collective action (and people who could have participated but did not), to observe the composition of their networks, and to have access to the relevant communications taking place within the network. We take advantage of a very rich dataset containing the Twitter networks (ties and ties-of-ties) for a group of participants and non-participants to the 2018 Women's March, as well as the communications related to the march that these users' networks exchanged the week before the protest. As previous studies show, Twitter networks are a good reflection of people's real-life networks (Bisbee and Larson, 2017), and the messages people are exposed to in social media are often responsible for the diffusion of protest movements online (Gonzalez-Bailon et al., 2011; Barbera et al., 2015; Casas and Webb Williams, 2018), indicating that social media communication offer a great opportunity to uncover different types of network effects (Jost et al., 2018).

We focus our analysis on four mechanisms scholars believe are responsible for the mobilizing role of personal ties. We call them the *information*, *motivation*, *coordination*, and *pressure* mechanism. We see them as *non-exclusive*, meaning that a single message can contain more than one of these mechanisms.

3.1 The Information Mechanism

Personal networks provide basic logistic information needed to participate. Information costs are a potential impediment for the success of collective actions (Olson, 1965). In order for people to decide whether to join a protest, they first need to be aware that such action is taking place, where and when, and how to get there. Communicating and consuming this logistic information has some costs attached to it.

One of the non-protesters in our datasets tweeted the following the day of the march: "*I'm kinda surprised by the #WomensMarch2018 photos and video from throughout the country. I admit I had a busy work week so I didn't consume much news, so I didn't even know about it. But clearly a lot of people did!.*"

This anecdotal evidence raises three main points. First, it illustrates the importance of exposure to basic logistic information. In a counterfactual scenario where this person had been exposed to information about where and when the march was taking place, she might have chosen to attend. Second, this example highlights the costs attached to consuming basic logistic information, and, in particular, that these costs are often low; our potential protester could have learned about the 2018 Women's March by watching the news on TV, reading a newspaper, or even seeing a Tweet. However, despite being low, not everybody is willing or able to bear them.

Finally, in this scenario one can also see how personal ties could have lowered the costs of obtaining the basic logistic information needed. This person could had easily learned about the march in a simple conversation with a friend who already possessed the information, or via a social media message from a friend in her network. A substantive number of existing studies stress that personal networks are mobilizing in part because they contribute to the transmission of necessary organizing information (Gould, 1991; Lohmann, 1994; Opp and Gern, 1993; Gonzalez-Bailon et al., 2011; Theocharis et al., 2015). Building on this literature, we formulate the following hypothesis:

Information($_{H_5}$): *Users whose friends sent Twitter messages containing basic logistic information related to the 2018 Women's March were more likely to protest.*

3.2 The Motivation Mechanism

Personal networks provide motivation to participate. Studies of opinion formation and political attitudes show that people often do no pay attention to politics (Hibbing and Theiss-Morse, 2002) nor do they have clear policy positions on a wide range of issues and policies (Converse, 2006). When it comes to making political decisions, such as voting and attending a protest, people often rely on cues they receive from their networks to form their own opinion. For example, Katz and Lazarsfeld (1955b) and Page and Shapiro (1992) find that media often influences public opinion by first influencing the views of attentive publics, who then influence the views of those in their networks.

Lohmann (1994) also argues that cascades of information providing reasons for people to protest was one of the main triggers of the demonstrations that took place in Leipzig in 1989-1991, previous to the collapse of the German Democratic Republic (GDR). In a context of total repression, where people knew very little about the performance of the regime, the initial demonstrations encouraged early core and informed activists to spread information about the 'nature of the regime' (Bikhchandani et al., 1992): "environmental damage, political repression, corruption, and the luxurious lifestyles of the SED elite" (Lohmann, 1994, 44) – cascades of motivation spread through personal networks and others decided to join the protests after learning what was actually going on in the country.

Some of the Women's March related messages that people sent on Twitter the days before the protest clearly provided motives for others to protest. The message shown in Figure 2 for example, apart from providing basic logistic information about the march (e.g. the day the march is taking place in Los Angeles and a link to further logistic information), it also

Figure 2: Example of a Tweet sent the week before the Women's March that had some of the theorized mechanisms.

lays out reasons for why people should care and attend the march. The text stresses the importance of supporting "diverse and powerful woman" and the images highlight some of the unequal conditions women face, such as only being paid on average 81% of the salary man receive for similar positions.

Hence, we build on the described literature on the mobilizing role of personal networks, opinion formation, and political participation to formulate the following hypothesis:

Motivation($_{H_6}$)**:** *Users whose friends sent Twitter messages providing reasons to attend to the 2018 Women's March were more likely to protest.*

3.3 The Coordination Mechanism

Personal networks solve coordination problems. One of the other main reasons for why collective actions often fail is because potential participants are uncertain about whether other people will join the action. This is particularly relevant, although not exclusive of, authoritarian regimes, where being the only one (or one of the few) to protest means being a clear repression target.

The literature on the role of personal networks in solving coordination problems became particularly salient after the unexpected collapse of the Soviet Union in the late 1980s. Kuran (1991) argues that people against the regime had clear incentives to falsify their preferences and to not show their private believes publicly until others around them had already done so. When people with a very high mobilization propensity started protesting on the streets, they 'infected' those in their personal networks with a lower propensity to participate by showing that more people were willing to protest, generating a 'bandwagon' effect that ended up with the collapse of the regime. Other models of network contagion corroborate Kuran (1991)'s work by also showing that personal networks can be mobilizing in part because they solve these type of coordination problems (Opp and Gern, 1993; Lohmann, 1994; Chwe, 2000).

In a democratic context like the United States, the work of others also show that the coordination role of networks can also play a mobilizing effect. In a study of the online diffusion of the Black Lives Matter movement, after controlling for a battery of alternative explanations, Casas and Webb Williams (2018) find that images shared online that showed people already protesting and supporting the movement on the street were in part responsible for an increase in attention to and diffusion of the movement online. What remains unclear is whether messages from friends that contribute solving coordination problems have an actual effect on real-world protest attendance. Apart from providing basic logistic information and reasons to protest, the message in Figure 2 also contributes to solving coordination problems by letting people know that the friend who sent the message will be attending the march in Los Angeles on January 20th. Hence, based on the rich literature on the matter, we formulate the following hypothesis:

Coordination($_{H_7}$): *Users whose friends sent Twitter messages stating that they would attend the 2018 Women's March, or that the march would be successful in terms of participation, were more likely to protest.*

3.4 The Pressure Mechanism

Personal networks put pressure to participate. Classic collective action theory predicts pressure mechanisms to increase the likelihood that an action will succeed (e.g. Olson (1965)'s "negative incentives"). A sense of social belonging, and ingroup-outgroup dynamics, explain a wide range of political attitudes: people do not desire to deviate from their group's behavior (Tajfel, 1981).

Gerber et al. (2008) work provides a clear example of how peer pressure can significantly influence political participation. In a study of voting behavior during the 2006 primary elections, the authors run a field experiment where a group of citizens received a letter (previous to the election) telling them that their neighbors would know whether they had voted or not in the upcoming election, putting peer pressure on them to vote.[6] The participation rate among these voters was about 10 percentage points higher than the average voter.

Building then on this existing literature on group dynamics and social pressure, we formulate the following hypothesis:

Pressure($_{H_8}$): *Users whose friends sent Twitter messages putting pressure on others to participate to the 2018 Women's March were more likely to protest.*

[6]Other randomized group of voters received other letters containing different treatments and placebo tests.

4 Data & Methods

We test our hypotheses by studying real-word attendance to the 2018 Women's March in the United States, and the Twitter networks and communications of a group of users who attended one of the marches and a group of users who showed support but did not attend. The first Women's March took place on January 21th 2017, right after Donald Trump's inauguration, with the goal of "harnessing the political power of diverse women and their communities to create transformative social change."[7] More than 650 demonstrations were organized all around the World, most of them in cities in the United States – about 400, with a principal march on Washington D.C..[8] Experts estimate that more than 4 million people participated to the U.S. protests alone, making it the largest demonstration in U.S. history.[9] On the anniversary of this first march, the organizers called for a second march in 379 U.S. locations (plus other locations around the World) on January 20th and 21st, with the epicenter being the march in Las Vegas.[10] Apart from pursuing the main mission of the organization, under the motto "Power to the Polls", this second march had as a particular goal to foster the election of political candidates all around the United States that supported gender equality and other movement demands.

Three types of data were needed in order to test the hypotheses proposed in this study: a) to find a group of people who protested and a group of potential protesters (users who

[7] Extracted form the Mission statement in the Women's March website.

[8] Number and location of the 2017 demonstrations obtained from the Women's March website. The link has now expired but an archived version can be accessed via the Web Archive in this url. We scrapped the data from the website before it disappeared, so we also have the information in dataset format.

[9] See the attendance data collected by Erica Chenoweth (University of Denver) and Jeremy Pressman (University of Connecticut) reported in this news article in Vox: *The Women's Marches may have been the largest demonstration in US History.*

[10] Whereas the main march in 2017 was the one in Washington D.C., in 2018 the organizers of the Women's March hosted a central event in Las Vegas, *Power to the Polls*, a kick-off event for a new national voter registration tour. Beside information about the march in Las Vegas, we also downloaded from the Women's March website a list of locations, and meta data about each of the other march that took place in the United States. As mentioned in Footnote 8, the information for the 2018 marches is no longer available in the website. We however have a dataset containing all the march data for replication.

positively messaged about the movement) that did not attend any of the marches, b) detailed information about these people's (Twitter) networks , and c) information about the communications and type of messages people exchanged through their networks via this social media platform the week before the protest. Before presenting our findings, we first describe in detail how we collected and preprocessed the data in order to have these three crucial pieces of information.

4.1 The Protester and Non-Protester Groups

To find a group of Twitter users who attended and a group who did not, we needed to collect all Twitter messages related to the march, look for users who geolocateed their tweets, and then check whether they sent a tweet from the march at the time the march closest to them was taking place – or whether they were in another location.

Before and during the march we collected Twitter messages mentioning a set of hashtags associated with the demonstrations using the Streaming API.[11] 15,679 users used a geolocated account (10,828 on January 20th and 4,851 on January 21st) – meaning that we had information about the place (city) from which these messages were sent (city and 4-coordinate-point bounding box around the locality).[12]

In order to know whether they sent a message at the time the march closest to them was taking place, and whether they sent it from the march location, we first needed to obtain

[11] We included in our search terms hashtags the organization asked supporters to promote as well as hashtags we saw (two weeks before the march) that users were using in relation to the demonstration: #womensmarch, #womensmarch2018, #powertothepolls, #togetherwerise, #whywemarch, #whyimarch, #imarchfor, #marchingforward, #womensunite, #unitedresistance, #resisttrump.

[12] In the past, geolocated Twitter messages used to reveal the specific coordinates from which a geolocated tweet was being sent (Larson et al., 2018), but after recent changes made by Twitter to increase user privacy, most of the time one can only access the name of the locality from which a given geolocated tweet is sent, and a 4-coordinate-point bounding box delimiting that locality. We use the information in the 'place' endpoint provided for each tweet by the Twitter API. Although this 'place' endpoint has a 'coordinates' field in it, for most tweets from geolocated accounts this field is empty, whereas the other fields providing information about the locality from which the user is sending the tweet (e.g. 'place_type', 'name', 'full_name', 'country_code', 'country', and 'bounding_box' around the location) always contain information.

detailed information about the location and time of each march. So the week following the march we scraped the Women's March website to collect information about all the marches that took place on January 20th (N=249) and January 21st (N=65), building a dataset with the followin information: day of the march, locality, start time, and address and coordinates of the starting point of each march.

We proceeded to match each user to a march by finding the march closest to any of the tweets sent by a user.[13] Then, as a first filter to find users for the protester and non-protester groups, we checked whether the 15,679 users with a geolocated account tweeted between the start time of the closest march and 8 hours after.[14] Out of the ones that did, we then considered the users who tweeted at least one message from the municipality where the closest march was taking place as a ***potential protester***, and the ones that did so from another place as ***potential non-protesters***.

At this point we were still uncertain about whether all the users in these two groups were actual protesters and non-protesters for three main reasons: a) some *actual* non-protesters could be in the group of *potential* protesters if they tweeted from the locality of their closest march but did not attend the march, b) some *actual* protesters could be in the group of *potential* non-protesters if they tweeted about their attendance to the march afterwards from another location, and c) some non potential participants could be in the *potential* protester and non-protester groups if they tweeted against the Women's March when the march closest to them was taking place, or if they were supportive of the Women's March but the closest march was too far away to feasibly attend the demonstration.

[13]By comparing the distances between the coordinates of each march starting point and the center of the 4-coordinate bounding box from each tweet.

[14]In this step we used a wide 8-hour time range for two main reasons. The first one is that some of the large marches lasted very long. The second is because the messages from users who tweeted from large marches, such as New York City, could not be actually sent during the march due to large concentration of Internet user in the same spot. These tweets were automatically sent hours later and the time-stamp of the tweet corresponds to this later time – we decided to go with a wider time window to avoid missing these protesters.

We used a three-fold strategy to address these issues and build the final protester and non protester set. First, we constrained the inclusion to the non-protester group to users who sent at least one message from 10 to 50 miles away from the closest protest, removing some non potential participants from the non-protester group (because they were too far away to attend). Second, one of the authors manually went through the messages the remaining users sent and checked whether they sent at least one message against the Women's March (see some examples in Appendix A): these non potential participants were also excluded from the dataset (these users were not of interest because they would had never attended a march no matter what). Finally, a research assistant went through the messages of the remaining users and indicated whether she thought they had attended a march, had not attended any, or she was not able to say based on the messages (3-class variable).[15] She also indicated whether the user was an organization such as a media outlet or a company (binary variable). The organizational accounts were removed from the dataset and we only kept in the *final* protester and non-protester groups those users that the research assistant also considered that had and had not protested, respectively. Finally, to simplify the rest of the data collection process and analysis, we focused next only on protesters who attended a march on January 20th (the day most of the marches took place) and the non-protesters who did not participate in a march in any of the two days (January 20th or 21st).

As a result, we ended up with two ***final groups*** of 2,607 and 507 people for whom we were highly confident that the person either attended or did not attend one of the 2018 Women's Marches, respectively.[16] These 3,114 users composed the final sample of our analysis.

[15] A second research assistant also coded about 30% of the messages to check for inter-code reliability. On these cases they both coded the two coders agreed about about 80% of the time.

[16] Having examined each of these accounts by hand, we are also confident that they are indeed people and not bots.

4.2 Protesters and Non-Protesters' Ties

The next step toward having all the necessary data to be able to test our hypotheses was to collect information about the Twitter networks of each of these 3,114 users.

Most works studying the mobilizing role of networks using real-world data tend to focus on people's immediate network (their friends) and do not take into account the effects that other further connections can have (such as friends of friends). An advantage of using social media data is that one can easily collect not only the list of user one follows (friends or ties) but also the list of followers for those followed by a given user (ties of ties). Moreover, as previously mentioned, Twitter networks have been found to be a fair representation of people's network on real life (Bisbee and Larson, 2017).

Hence, the week after the weekend the Women's Marches took place, we collected the lists of people these 3,114 users followed (a total of 3,417,056 friends, 1,642,218 of them unique), and the lists of users these 1,642,218 unique friends followed (about 17 billion *non-unique* users).[17] This allowed us to build a comprehensive graph of the networks of the protesters and non-protesters in our sample.

4.3 Pre-March Network Communications

As we mentioned, during the week before the Women's March and while the demonstrations were taking place, we collected all Twitter messages mentioning the key hashtags related to the protest.[18] After collecting the list of friends of the users in the protester and non-protester sent, we went back to the message collection and pulled all the march-related messages that each user's friends sent during the week before the march. Two sets of messages were ignored: a) messages against the march (see Appendix A for some examples), and b) messages related

[17] Due to computation restrictions, we have not yet computed the exact number of friends of friends involved in the study, nor the number of these that are unique. The 17 billion is an estimate based on the average number of ties of ties the users in our sample have: see Figure 3

[18] See Footnote 11 for a list of the hashtags.

to a pro-life protest (that took place the day before, on January 19th) that shared a couple of the mobilizing hashtags with the Women's March (#WhyIMarch and #WhyWeMarch). Using the machine learning techniques that we will discuss next, these messages were detected and removed. We ended up with a final dataset of pre-march communications containing a total of 486,007 messages, 45,717 of them unique.[19]

Table 1: Example tweets for each of the theorized mechanisms.

Information
On 1/21, @womensmarch kicks off a year-long #PowerToThePolls campaign to win in 2018! The rise of the woman IS the rise of the nation. In 2018, lets rise together! https://t.co/O9sqO9ZXDU
The speakers lineup has been released for the #PowerToThePolls #WomensMarch in Huntsville this Saturday, 20 January, 11 am! #p2phsv https://t.co/I7okJ7KwVF
If you plan to ride any bus line being diverted to Mission Street, reconsider your options. Taking the 6-Parnassus was slow and wait times was 40+ minutes. This also affects Golden Gate Transit that regularly runs on Mission. #WomensMarch #SFMUni #SFMTA #Muni #GoldenGateTransit https://t.co/nSHgzrbsPA
Motivation
RT @ACLU_Mass: Tomorrow, we are taking to the streets for #WomensMarch2018 and organizing to fight Trump's anti-immigrant agenda.
Why would any citizen not register and vote? People died to give us this right! C'mon people! https://t.co/oXpNClY5TK #PowerToThePolls
History was made when more than 5M people across the world rallied in the #WomensMarch for equality. Here is some of last year's best moments that launched us towards progressive strides for equality from #MeToo & #TimesUp, to making bids for public office in the thousands. https://t.co/8Tkn30Z6z6
Coordination
What should my sign say for @womensmarch DC? #WomensMarch2018 dc
Front and back of signage. Carried it last year and carrying it again this year. #WomensMarch @womensmarch_sd @womensMarch #WomensMarchSD https://t.co/VaesZT3JZf
Marching with my 75-year-old mom! #WomensMarchNYC https://t.co/Pr2RBBZkIw
Pressure
@AoDespair "Liberty, Equality, Justice" is the motto of #AmericanVelvetRevolution we must start the #AmericanVelvetRevolution now! to the streets... join with #WomensMarch2018 in two days, do not obey dictates of #TrumpFascism, #VoteBlue every chance you get from local on up!
Are you marching Jan. 20? #WomensMarch https://t.co/7w5NXKt1t8
We all need to take to the streets tomorrow. Not just women. ALL Resistors. #resist #TrumpShutdown #womensmarch2018 Cleveland

[19]In the 486,007 message count we count a friend's message as many times as users in our sample follow her.

The next step was to detect the presence – or lack thereof – of each of the four mobilizing mechanisms in these tweets. First, we draw a random sample of 3,300 messages for two manual coders to label. Each manual coder labeled 1,850 tweets, with a set of 400 they both coded (see codebook in Appendix B). 200 of the 400 were coded multiple times and were used for training and calibration. The other 200 were used to calculate inter-rater reliability (IRR) measures (see Appendix C for detailed IRR statistics), which were satisfactory. Table 1 shows some example tweets coded as containing each of the theorized mechanisms.

Then we used this labeled dataset to train machine learning classifiers to predict six tweet-level outcomes: whether the tweet was against the march (sentiment), whether it was related to the pro-life mobilization that took place the day before (prolife), as well as whether the messages contained each of the mobilizing mechanisms (information, motivation, coordination and pressure). We trained an ensemble of six machine learning models using different types of tweet-level features as input: unigrams from the tweet text, unigrams from the text that shows up in tweet as a result of sharing a link, text embeddings, image embeddings, and joint text and image embeddings (see Appendix D for a detailed description of the models).

Table 2: Accuracy of the machine learning classifiers predicting the 4 mobilizing mechanisms (Information, Motivation, Coordination, and Pressure) and whether messages are against the Women's March (Sentiment) and about the Pro-Life protest that took place the day before the march (Prolife).

	Accuracy	F-Score	Class Prop.	Precision	Recall
Information	0.83	0.80	0.38	0.75	0.86
Motivation	0.77	0.67	0.35	0.71	0.64
Coordination	0.82	0.58	0.20	0.54	0.62
Pressure	0.81	0.51	0.20	0.55	0.47
Prolife	0.98	0.91	0.09	0.83	1.00
Sentiment	0.97	0.98	0.96	0.97	1.00

We assessed the ability of the six machine learning classifiers to predict each of the six binary outcomes, and for each outcome we selected the most accurate one. In Table 2 we report the final accuracy of the machine learning models used to predict the presence of each outcome. The overall accuracy is very high for all models (about 80% or higher). For the information, motivation, prolife and sentiment classifiers, precision and recall is also high (between 60 and 100%). The coordination and pressure classifiers are a bit nosier, precision and recall is about 50%. However, these are still substantively accurate classifiers: these are infrequent classes and by random chance we would correctly predict a tweet with these two mechanisms only about 20% of the time (see *Class Prop.* column in Table 2). As a final steps, we used these machine learning classifiers to detect and remove the messages sent by friends of protesters and non protesters that were against the Women's March and/or about the Prolife demonstration, and then we used the other four models to predict the presence of the four mobilizing mechanisms in the remaining 45,717 unique messages.

5 Results

We begin by testing the first four hypotheses and corroborating that users in the protester group were indeed more connected among themselves than were users in the non-protester group. We then transition to explore the Twitter communications of the protesters and non-protesters' networks in order to address the rest of the hypotheses. We first look at the number of messages related to the Women's March that Twitter friends from users in both groups sent the week before the protest. We then explore the presence in these messages of the four mechanisms of interest, and we conclude by estimating their effect on protest mobilization when controlling for other relevant covariates. In the final section we will discuss the implications of the findings as well as potential steps for future research.

In Table 3 we provide a summary of the network attributes of 507 non-protesters and a random sample of 507 of the 2,607 protesters, as well as group comparisons (t-tests) for each of the attributes. The goal of these comparisons is to evaluate the extent to which protesters were more connected among themselves than non protesters. If personal networks have no effect on people's decision to protest (null hypothesis for $H_{1,2,3,4}$), we should see no differences in group connectivity. To make sure the results were not a simple function of group size, we decided to draw and evaluate a random sample of protesters that was equal in size to the non protester set (n = 507). Positive t-statistics indicate a larger attribute average for the protester group whereas negative t-statistics indicate the opposite.

Two main points stand out from Table 3. First, non protesters had larger networks than users in the protester group. On average, they had a greater number of ties (1,244 *versus* 1,073 for protesters), ties of ties (about 6.4 million *versus* 4.6 million for protesters), and reciprocated ties (502 *versus* 419 for protesters). Moreover, non-protesters had a reciprocated friendship with a larger proportion of their friends (.32 *versus* .31 for protesters) and they had a larger proportion of triadic relationships –out of all possible triadic connections with their friends and friends of friends– (.11 *versus* .1 for protesters). The group differences for the number of ties of ties and the proportion of triadic connections (rows 3 and 5 in Table 3) are statistically significant at conventional levels.

Table 3: Network attributes for 507 Protesters and 507 Non-Protesters (with standard deviation in parentheses). *T-statistic* tests the null hypothesis that attributes for Protester and the Non-Protesters are the same.

	PROTESTERS	**NON-PROTESTERS**	**tstat**
# Users in each set	507	507	–
Mean # Ties	1,073 (1,903)	1,244 (1,837)	-1.45
Mean # Ties of Ties	4,618,019 (5,818,521)	6,437,016 (7,689,719)	-4.23
Mean # Reciprocated	419 (1,370)	502 (1,028)	-1.09
Mean Prop Reciprocated	0.31472 (0.20377)	0.32061 (0.20799)	-0.45
Mean Transitivity	0.09789 (0.05982)	0.10676 (0.0797)	-1.99
Prop. Ties Within ($\mathbf{H_1}$)	0.00033 (0.00089)	0.00012 (0.00046)	4.70
Prop. Ties of Ties Within ($\mathbf{H_2}$)	0.00005 (0.00008)	0.00004 (0.00003)	2.19
Prop. Reciprocated Within ($\mathbf{H_3}$)	0.00047 (0.00308)	0.00013 (0.00072)	2.37
Prop. Triadic Ties Within ($\mathbf{H_4}$)	0.00010 (0.00038)	0.00004 (0.00016)	3.39

However, as the information at the bottom half of Table 3 indicates (rows in gray), despite having smaller networks, the Twitter users that attended one of the Women's March on January 20th were indeed more connected among themselves. A larger proportion of their ties was also present in the protester group (.0003 *versus* .0001 for non-protesters) as well as a greater proportion of their ties of ties (.00005 *versus* .00004 for non-protesters), reciprocated friends (.0005 *versus* .0001 for non-protesters), and triadic ties (.0001 *versus* .00004 for non-protesters). In line with existing network models of protest mobilization (Granovetter, 1978; Marwell et al., 1988; Siegel, 2008; Centola and Macy, 2007; Larson et al., 2018), all these positive and statistically significantly differences in favor of the protester group corroborate the first four hypotheses of the study ($\mathbf{H}_{1,2,3,4}$).

However, what explains these differences? Why are personal networks mobilizing? Can we gain some leverage by exploring the communications that the protester and non-protester networks exchanged the week before the protest? Figure 3 shows the average daily original tweets (on the left) and retweets (on the right) about the Women's March that Twitter friends of users in the two groups sent from January 15th to January 19th (both included). Figure 4 shows the median instead of the average number of messages. As one would expect, the amount of discussion about the march increases as the demonstration is closer in time, particularly during the two days preceding the demonstrations (January 18th and 19th). On average, during these preceding days, users in the non protester set were exposed to more messages than the users in the protester group. This can easily be a function of non protesters having a larger number of friends and so having higher initial chances to be exposed to Women's March messages. However, Figure 4 reveals that, despite having a smaller average number of friends, users in the protester set saw the same median number of march-related messages; and on January 19th the median user in the protester set actually saw more original march-related messages than the median user in the non protester set.

Figure 3: Average Daily Women's March related messages the Twitter friends of users in the Protester and Non-Protester sent during the 5 days previous to the protest.

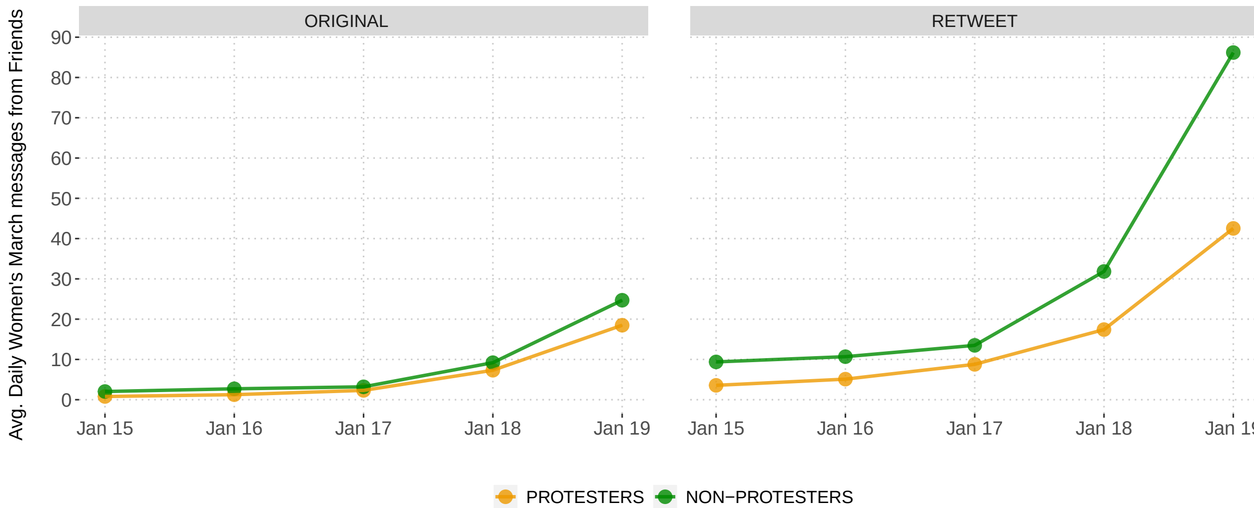
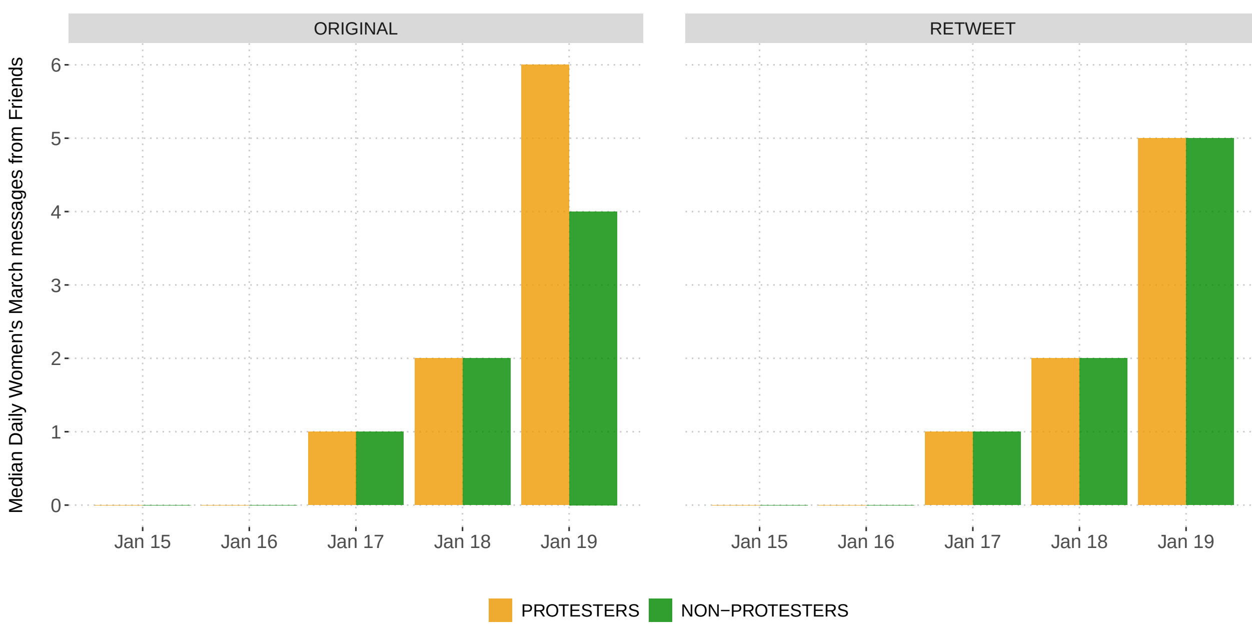
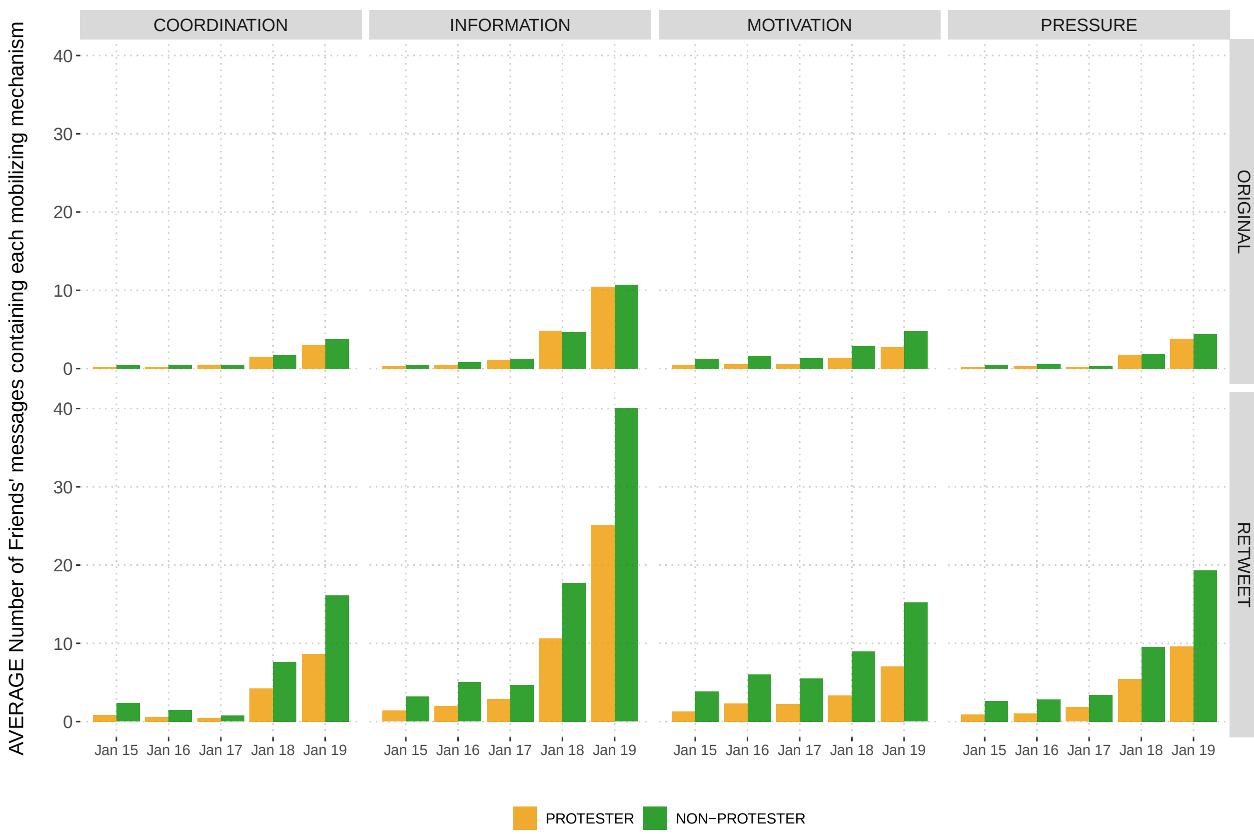


Figure 4: Median Daily Women's March related messages the Twitter friends of users in the Protester and Non-Protester sent during the 5 days previous to the protest.



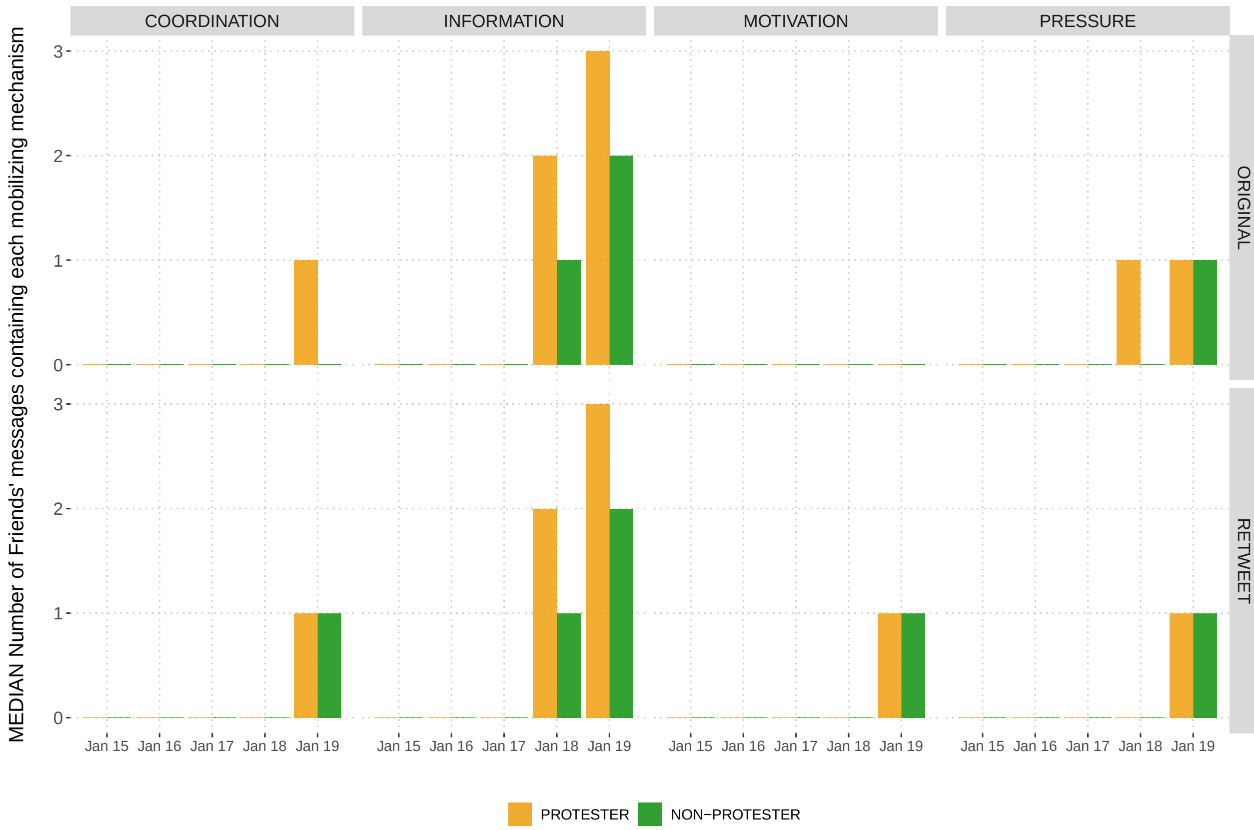
What remains unanswered however is whether the discussion among the networks of the users in the protester and non-protester set was similar in content. Did the Twitter friends of the protesters used the theorized mobilizing mechanisms at a higher rate?

Figure 5: Average number of Twitter messages to which users in the protesters and non protester set were exposed the week before the Women's March.



To address this question we used the trained machine learning models to predict which of the tweets sent by the protesters and non protesters' friends contained each of the mobilizing mechanisms. In Figure 5 we see that the average user in the non protester set was actually exposed to a larger number of original tweets, but particularly retweets, containing each of the mechanism. In Figure 6 we show median instead of average values and observe that the

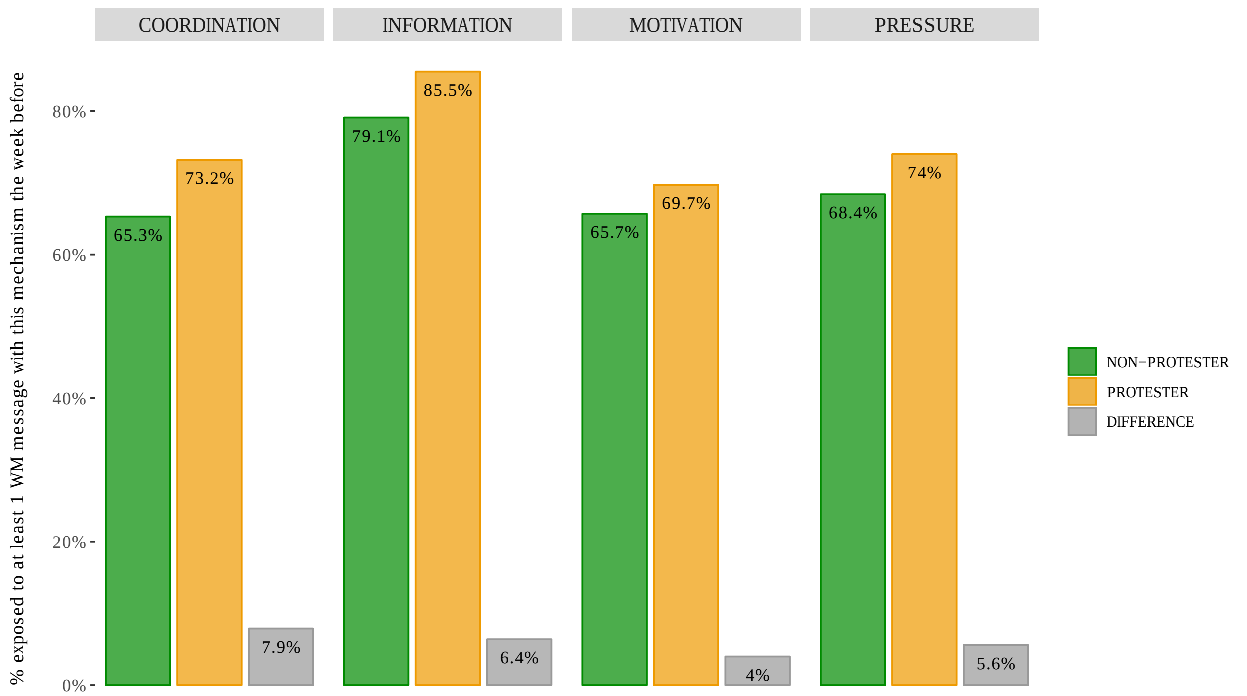
Figure 6: Median number of Twitter messages to which users in the protesters and non protester set were exposed the week before the Women's March.



results in Figure 5 are driven by some outliers: the median protester was actually exposed to a larger number or original tweets and retweets containing the mechanisms. In particular, a larger number of messages containing the coordination and information mechanism.

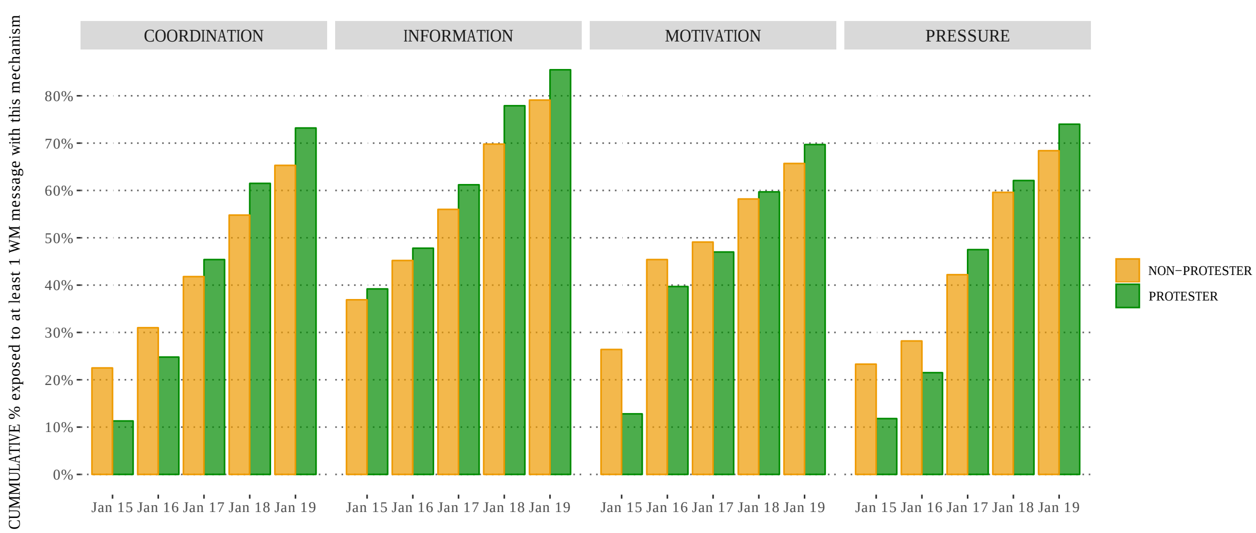
A problem with focusing the analysis on average and median group values is that we cannot get a clear sense of the amount of users in each set that were actually exposed to these messages. For this reason in Figure 7 we show the percentage of protesters and non protesters who saw at least one message containing each of the mechanisms. We clearly observe that a larger proportion of users in the protester set were exposed to these type of messages. As we saw in Figure 7, we observe the largest differences when we compare the percentage of users exposed to the coordination and information mechanism: about 8 and

Figure 7: Percentage of users in the protester and non protester groups who were exposed to at least one message containing each of the mobilizing mechanisms



6.5 percentage points difference in favor of the protesters.

Figure 8: Cumulative percentage of users in the protester and non protester groups who were exposed to at least one message containing each of the mobilizing mechanisms during the week before the Women's March.



In Figure 8 we show the moment in which users in both sets first saw a message containing each mechanism the week before the march. These are cumulative percentages, meaning that users counted in the percentages from an early day are only counted once and carried on to the next percentage. For three of the mechanisms (coordination, motivation, and pressure), we actually see that four days before the protest a larger percentage of non protesters had been exposed to at least one message containing those mechanisms. However, as the day of the marches approached, those percentages flipped to become the positive rates in favor of the protesters that we reported in Figure 7.

Table 4: Logistic regression predicting protest attendance.

	β	SE
(Intercept)	3.3677	0.3934*
Number of Friends (log)	-0.3949	0.0667*
Num. WM messages sent by friends	0	0.0018
Num. unique friends who messaged about the WM	0.0032	0.0037
Num. unique reciprocal friends who messaged about the WM	0.0003	0.0027
Num. INFORMATION messages sent by friends	0.0136	0.0064*
Num. MOTIVATION messages sent by friends	-0.0031	0.0029
Num. COORDINATION messages sent by friends	-0.0182	0.0099
Num. PRESSURE messages sent by friends	-0.0124	0.0069
Exposure to at least 1 INFORMATION message	0.4636	0.1932*
Exposure to at least 1 MOTIVATION message	0.0028	0.1617
Exposure to at least 1 COORDINATION message	0.4265	0.1784*
Exposure to at least 1 PRESSURE message	0.0261	0.1961
Gender Unclear [Ref = Male]	-0.8263	0.2492*
Gender Female [Ref = Male]	0.0502	0.1127
Political interest (Num. members of Congress followed)	-0.0188	0.005*
Political interest (Num. top-50 media accounts followed)	0.0151	0.0124
Interest in the movement (follows WM account)	-0.009	0.3389
Conservatism (following FoxNews)	-0.7705	0.1914*
Liberalism (following more Dem. than Rep.)	0.0328	0.0101*
N = 3,056		
Log Likelihood = −1,260.720		
AIC = 2,561.441		

In Table 4 we evaluate the extent to which these group differences hold when controlling for other confounders. As we also observed in Figures 6 and 7, there is a clear relationship between exposure to information and coordination type messages and attendance to the 2018 Women's March. The larger the number of messages with basic logistical information to which people were exposed, the larger the likelihood they attended the march. Moreover, exposure to at least one information message, and exposure to at least one message reporting that a friend was attending the march, was also strongly related to higher attendance. Overall, in the previous figure but particularly in Table 4, we find strong evidence supporting the hypotheses that personal networks are in part mobilizing because they provide basic logistic information needed for protest attendance (H_5) and because they contribute to solving coordination problems (H_7). However, although we saw in Figure 6 and 7 that a larger proportion of protesters (compared to non protesters) saw at least one motivation and pressure type message, once we control for other covariates we do not find enough support for H_6 and H_8.

6 Discussion

The argument that personal ties have an effect on people's decision to protest is a longstanding one in a wide range of social science disciplines (Granovetter, 1973, 1978; Marwell et al., 1988; Siegel, 2008; Centola and Macy, 2007; Larson et al., 2018). Nevertheless, the reasons why personal networks play such a mobilizing role are much less established. A more clear picture of the network-mobilizing effect is of great importance if we are to understand the conditions under which protest movements emerge, as well as the condition under which social groups can set political agendas and influence policy.

In this paper we shed new light onto this question by studying the networks and the network communications of a group of Twitter users who attended one of the Women's

March in the United States on January 20th 2018, and a group of users who showed interest by messaging about the march but did not attend.

We tested four hypotheses regarding the mechanism by which exposure to tweets through one's social networks could impact an individual protest decision: providing basic logistical *information* needed to protest, creating *motivations* for others to protest, solving *coordination* problems by revealing protest attendance, and receiving *pressure* from others to attend. We used machine learning techniques to measure the presence of these different types of messages in the tweets to which protesters and non-protesters sent the week before the march.

Our findings were extremely interesting. Our network based analysis largely matched the findings from (Larson et al., 2018), only this time in a different country and with a different type of protest: users in the protester group were more connected among themselves than were users in the non-protester group. Then, when studying the communications these users' networks exchanged the week preceding the protest, we found that exposure to messages containing basic logistic information needed for protest attendance, and messages communicating that other friends would attend, increased users' likelihood to attend. However, we did not find enough evidence in support of the argument that personal networks mobilize because they provide motives or put pressure on others to attend. This null effect does not necessarily mean that personal networks are not mobilizing because of these two reasons. It could also be that these other types of influence are mainly happening offline.

In conclusion, this paper not only provides a very clear picture of the conditions under which personal networks influence one's decision to protest, but it also opens the door to future research on the matter. The methods applied in this paper can be used to study many other protests, particularly mobilizations that have been planned ahead of time, allowing the researcher to start tracking the messages potential protesters exchange the days previous to the action.

References

Barbera, Pablo, Ning Wang, Richard Bonneau, John T. Jost, Joshua Tucker, and Sandra Gonzalez-Bailon. The critical periphery in the growth of social protests. Plos One, 2015.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch. A theory of fads, fashion, custom, and cultural change as informational cascades. Journal of Political Economy, 100(5): 992–1026, 1992.

Bisbee, James and Jennifer M. Larson. Testing social science network theories with online network data: An evaluation of external validity. American Political Science Review, 111 (3):502–521, 2017.

Casas, Andreu and Nora. Webb Williams. Images that matter: Online protests and the mobilizing role of pictures. Political Research Quarterly, xx(s):xx–xx, 2018.

Centola, Damon and Michael Macy. Complex contagions and the weakness of long ties. American Journal of Sociology, 113(3):702–734, 2007.

Centola, Damon M. Homophily, networks, and critical mass: Solving the start-up problem in large group collective action. Rationality and Society, 25(1):3–40, 2013.

Chwe, Michael Suk-Young. Communication and coordination in social networks. The Review of Economic Studies, 67(1):1–16, 2000.

Converse, Philip E. The nature of belief systems in mass publics. Critical Review, 18(1-3): 1–74, 2006.

Frostenson, Sarah. The women's marches may have been the largest demonstration in u.s. history. Vox, Jan 2017.

Gerber, Alan S., Donald P. Green, and Christopher W. Larimer. Social pressure and voter turnout: Evidence from a large-scale field experiment. American Political Science Review, 102(1):33–48, 2008.

Gonzalez-Bailon, Sandra, Javier Borge-Holthoefer, Alejandro Rivero, and Yamir Moreno. The dynamics of protest recruitment through an online network. Scientific Reports, (197), 2011.

Gould, Roger V. Multiple networks and mobilization in the paris commune, 1871. American Sociological Review, 56(6):716–729, 1991.

Granovetter, Mark. Threshold models of collective behavior. American Journal of Sociology, 83(6):1420–1443, 1978.

Granovetter, Mark S. The strength of weak ties. American Journal of Sociology, 78(6): 1360–1380, 1973.

Hibbing, John R. and Elizabeth Theiss-Morse. Stealth Democracy: Americans' Beliefs About How Governmnet Should Work. Cambridge University Press, Cambridge, UK, 2002.

Jost, John T., Pablo Barbera, Richard Bonneau, Melanie Langer, Megan Metzger, Jonathan Nagler, Joanna Sterling, and Joshua A. Tucker. How social media facilitates political protest: Information, motivation, and social networks. Political Psychology, 39(S1):85–118, 2018.

Katz, E. and P.F. Lazarsfeld. Personal Influence, the Part Played by People in the Flow of Mass Communications. Free Press, 1955a.

Katz, Elihu and Paul F. Lazarsfeld. Personal Influence: The Party Played by People in the Flow of Mass Communications. Free Press, New York, 1955b.

Kuran, T. Private Truths, Public Lies: The Social Consequences of Preference Falsification. Harvard University Press, 1995.

Kuran, Timur. Now out of never: The element of surprise in the east european revolution of 1989. World Politics, 44(1):7–48, 1991.

Langer, Melanie, John Jost, Richard Bonneau, Megan MacDuffee Metzger, Sharareh Noorbaloochi, and Duncan Penfold-Brown. Digital dissent: An analysis of the motivational contents of tweets from an occupy wall street demonstration. Motivation Science, 02 2018.

Larson, Jennifer M., Jonathan Nagler, Jonathan Ronen, and Josh Tucker. Social networks and protest participation: Evidence from 130 million twitter users. Working Paper, x(x): xxx–xxx, 2018.

Lohmann, Susanne. The dynamics of informational cascades: The monday demonstrations in leipzig, east germany, 1989–91. World Politics, 47(1):42–101, 1994.

Marwell, Gerald, Pamela E. Oliver, and Ralph Prahl. Social networks and collective action: A theory of the critical mass. iii. American Journal of Sociology, 94(3):502–534, 1988.

McAdam, Doug and Ronnelle Paulsen. Specifying the relationship between social ties and activism. American Journal of Sociology, 99(3):640–667, 1993.

Olson, Mancur. The Logic of Collective Action: Public Goods and the Theory of Goups. Harvard University Press, Cambridge, MA, 1965.

Opp, Karl-Dieter and Christiane Gern. Dissident groups, personal networks, and spontaneous cooperation: The east german revolution of 1989. American Sociological Review, 58(5):659–680, 1993.

Page, Benjamin I. and Robert Y. Shapiro. The Rational Public: Fifty Years of Trends in American's Policy Preferences. The University of Chicago Press, Chicago, 1992.

Siegel, David A. Social networks and collective action. American Journal of Political Science, 53(1):122–138, 2008.

Tajfel, Henri. Human Groups and Social Categories: Studies in Social Psychology. Cambridge University Press, Cambridge; New York, 1981.

Theocharis, Yannis, Will Lowe, Jan W. van Deth, and Gema Garcia-Albacete. Using twitter to mobilize protest action: online mobilization patterns and action repertoires in the occupy wall street, indignados, and aganaktismenoi movements. Information, Communication & Society, 18(2):202–220, 2015.

Appendix A Examples of Anti Women's March Tweets

Table 5: Examples of Anti Women's March Tweets

Hello everyone! We like all humans except white straight males ... and women who dont fans of abortion #WomensMarch2018
#WomenMarch2018 #WomensMarchNYC This movement supports a culture that will RAPE YOU, Stone you, Cover you, and take your car keys, and if that doesnt work, honor kill you, I apologize for mansplaining. https://t.co/VaQFRDXmL4
Women's march not concentrating on REAL women's issues.,Like being victims of crime.,From both illegal and legal immigrants.,Look at the statistics.,What about 2nd amendment rights for women?,Physically a woman can't beat a man, but a 230 45 ACP bullet CAN! #WomensMarch2018
I can not take this! I AM NOT THIS KIND OF WOMAN. #WomensMarch2018RT if you're NOT THIS KIND OF WOMAN! https://t.co/MW3DaZmnwD
This march is, in general, for misandrists who have made themselves perpetual victims of having a vagina. Freud called it Penis Envy. #WomensMarch2018
Wow! How can I un-see these horrific pics. Democrats out today. Our ugly part of America out today! https://t.co/xaDa39mMN7
Two years in a row and the activists at the #WomensMarch have failed to answer one fundamental question. What specific rights do you not have an equality of opportunity towards?
I'll ask the same question I asked last year: what the hell are these crazy Leftist feminists marching for? Last time I checked women of America are free and liberated! #WomensMarch2018 #WomensMarchDC #SaturdayMorning https://t.co/D162HxLH1k
I support #MarchforLife not this made up Liberal #WomensMarch2018 https://t.co/0ClFFaL4QC

Appendix B Codebook for Coding Pre-March Tweets from Protesters and Non-Protesters' Friends

Purpose: Coding pre-march tweets for the type information users in our protester and non-proster were exposed few days before the march.

General Instructions:

- Each coder should have access to a Google spreadsheet where they will insert the labels for each message, and also to a folder with html files showing messages in a Twitter-looking format.
- Coders need to click on the html files, judge whether messages in there contain one of the 4 mobilizing mechanisms. [If the message has a link, coders DO NOT need to click on the link and look at the information in the link when judging the presence of the different mechanisms]
- None of these variables are mutually exclusive! So you should consider each variable for each message

---

Specific Instructions:

VAR 1 [information] Information Mechanism

Logic: In order to decide whether to attend a protest, people first need to have basic logistic information about where and when the protest is taking place, how to get there, etc. Conditional on ones initial predisposition to protest, one should be more likely to do so when provided with such basic information.

Instructions for coders: Is the tweet providing basic practical information about how to participate in a Womens March, such as the time of the march, its route, how to get there, events taking place before-during-after the march, etc?
[label = 1 if this mechanism is present. 0 or blank otherwise]

VAR 2 [motivation] Motivation Mechanism

Logic: Most people do not pay attention to politics and they lack clear policy positions. However, the average citizen often takes cues from more attentive publics when making political decisions. Those who we follow in social media can increase our likelihood of attending a protest by presenting us clear reasons for why we should do so.

Instructions for coders: Is the tweet providing a reason for attending a Womens March? This includes reasons to dislike Trump, Trump administration, Trump policies, and especially Trump policies/comments/preferences towards women. Motivation we are going to be open to reasons that might motivate one to protest, even if they are not specifically about the march.
[label = 1 if this mechanism is present. 0 or blank otherwise]

VAR 3 [pressure] Social Pressure Mechanism

Logic: A sense of social belonging, and ingroup-outgroup dynamics, explain a wide range of political attitudes: people do not desire to deviate from their groups behavior. Those who we follow in social media may send us messages clearly signaling that our network will not appreciate-tolerate inaction; pressuring us to protest.

Instructions for coders: Do you believe that this message is putting pressure on others to attend the Womens March? This includes message instructing people to go: e.g. You should go!
[label = 1 if this mechanism is present. 0 or blank otherwise]

VAR 4 [coordination] Coordination Mechanism

Logic: individuals deciding whether to attend a protest often face a coordination problem. If a large group of people protest, the demonstration is a success and one is not wasting the time by attending. However, since it is hard to know ex ante who and how many people will attend, one may decide to not bare the costs of attending and stay home.

Instructions for coders: Is someone in this message indicating their attendance to the upcoming Womens March? And/or is the message indicating that there will be a high attendance to the March?
[label = 1 if this mechanism is present. 0 or blank otherwise]

VAR 5 [prolife] ProLife-only messages

Logic: some of the hashtags we used to collect Womens March messages were also being used in a ProLife march that took place the day before: “whywemarch” and “whyimarch” mostly. I filtered out messages mentioning: whywemarch, marchforlife, prolife, whyimarch, prochoice, life, unbornlivesmatter, lovesaveslives; AND NOT mentioning “womensmarch”, but some ProLife messages may still be present in this dataset of pre-march messages to code. We want to detect them.

Instructions for coders: Is the message ONLY related to ProLife policy positions, demonstration, action; and not related at all to the Womens March?
[**label = 1** if only has a ProLife component and its not related to the Womens March, **label = 0** if its about the Womens March (even if it also has a ProLife component), **label = 9** if it’s unclear]

VAR 6 [sentiment] Sentiment Category

Instructions for coders: Do you think the message is in support of or against the Womens March?
[**label = 1** if the tweet is in favor of the Womens March, **label = -1** if the tweet goes against the Womens March, **label = 9** if neutral tone or unclear about the Womens March (for example, tweets against trump that dont directly mention the Womens March, or ProLife tweets that arent necessarily against the Womens March. If a quote is in favor but the quoted message against –or vice versa–).]

Appendix C Inter Rater Reliability

In this Appendix we report information to assess the *inter rater reliability* (IRR) of the two manual coders that manually labeled the tweets for whether they had each of the 4 mobilizing mechanisms (information, coordination, motivation, and pressure), whether they were related to the pro-life organization that took place the day before, on January 19th, and whether they were against the women's march (sentiment).

We draw a random sample of 3,300 messages sent by friends of users in the protester and non protester set. Then each manual coder labeled 1,850 tweets, with a set of 400 they both coded (see codebook in Appendix B). 200 of the 400 were coded multiple times and were used for training and calibration. The other 200 were used for IRR assessment.

Figure 9: Confusion Matrices for 2 manual coders labeling tweets for the 4 mobilizing mechanisms (coordination, information, motivation, and pressure) and 2 other variables of interest (prolife, sentiment)

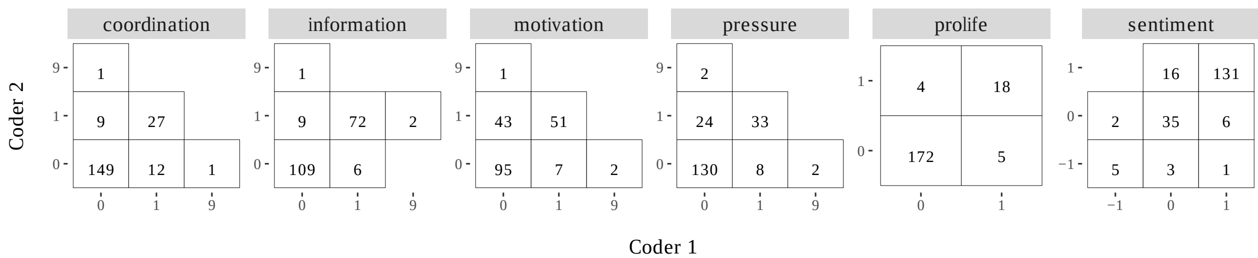


Figure 9 shows confusion matrices illustrating the agreement among the two coders. For the four mobilizing mechanisms, a 0 indicates the mechanism is not present, a 1 indicates the mechanism is present, and a 9 indicates that is unclear. For the prolife variable, a 0 indicates the tweet is about the Women's March whereas a 1 means that it is about the Pro-Life demonstration that took place the day before. For the sentiment variable, a -1 means the tweet is against the Women's March, a 0 indicates that it is about the march but has a neutral tone, and a 1 means the tweet is clearly in favor of the march (we collapsed the 0s and the 1s for this last variable when train the binary classifier predicting messages against the march).

In Figure 9 we observe very high agreement for 2 of the 4 mobilizing mechanisms (coordination and information) and the prolife and sentiment coding, and medium agreement for the other 2 mobilizing mechanisms (motivation and pressure). The lower agreement for these two mechanisms is the result of Coder 1 having a slightly more restrictive understanding of what constitutes motivation and pressure.

In Figure 10 we show the agreement rates for the 4 mobilizing mechanisms. A *flexible* agreement means that we consider as an agreement when a coder said that it was unclear whether a mechanism was present and the other coder provided a 0 or 1 label. A *strict* agreement means that we only consider the coders agree when both labeled the tweet as a 0 or as a 1. Overall, we observe the agreement to be higher than 80%. The only exception is the motivation mechanism, for which agreement was around 75%.

Figure 10: Coding agreement for 2 coders labeling 200 tweets for the presence of the 4 mobilizing mechanisms.

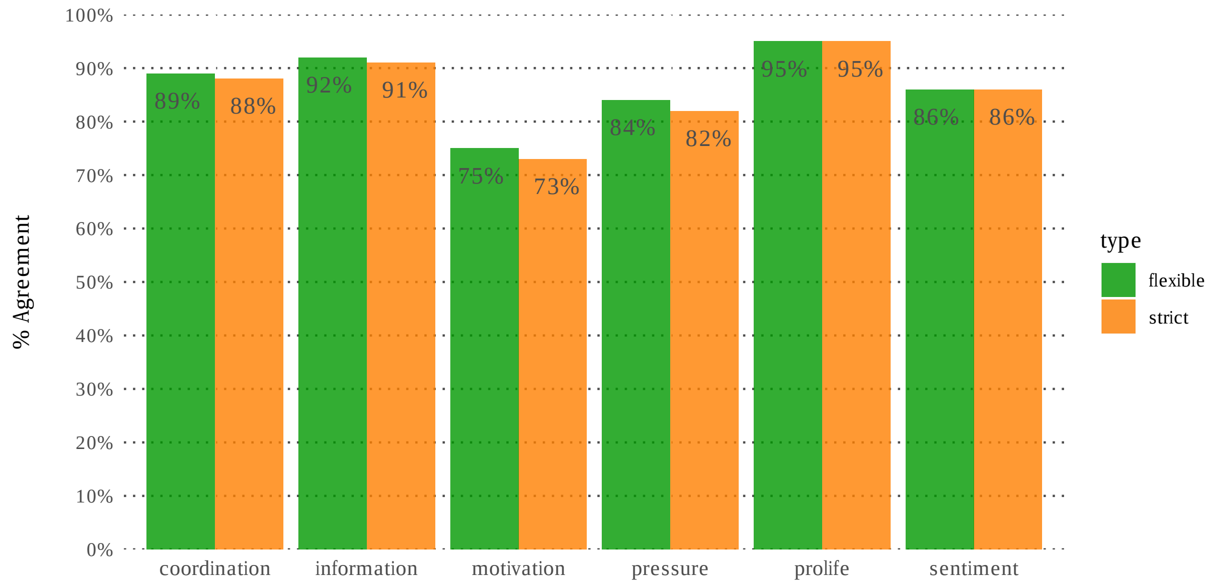


Figure 11: Inter Rater Reliability statistics for 2 coders labeling 200 tweets for the presence of 4 mobilizing mechanisms.

	Prop. Agreement	Krippendorff's Alpha	Cohen's Kappa	Scott's Pi
information	0.92	0.84	0.84	0.84
motivation	0.74	0.46	0.48	0.46
pressure	0.84	0.56	0.57	0.56
coordination	0.89	0.66	0.65	0.65
prolife	0.95	0.78	0.77	0.77
sentiment	0.88	0.68	0.68	0.68

Note: A statistic between 0 and .20 indicates *poor* agreement, .21-.40 *fair* agreement, .41-.60 *moderate* agreement, .61-.80 *substantial* agreement, and .81-1.00 *near perfect* agreement.

Table 11 reports common IRR statistics (Krippendorff's Alpha, Cohen's Kappa, and Scott's Pi) for the 6 variables labeled by the manual coders. We observe that in all cases the IRR statistics are above .4 (indicating at least acceptable *moderate* agreement), but in most cases we observe *substantial* (>.6) or *near perfect* agreement (>.8).

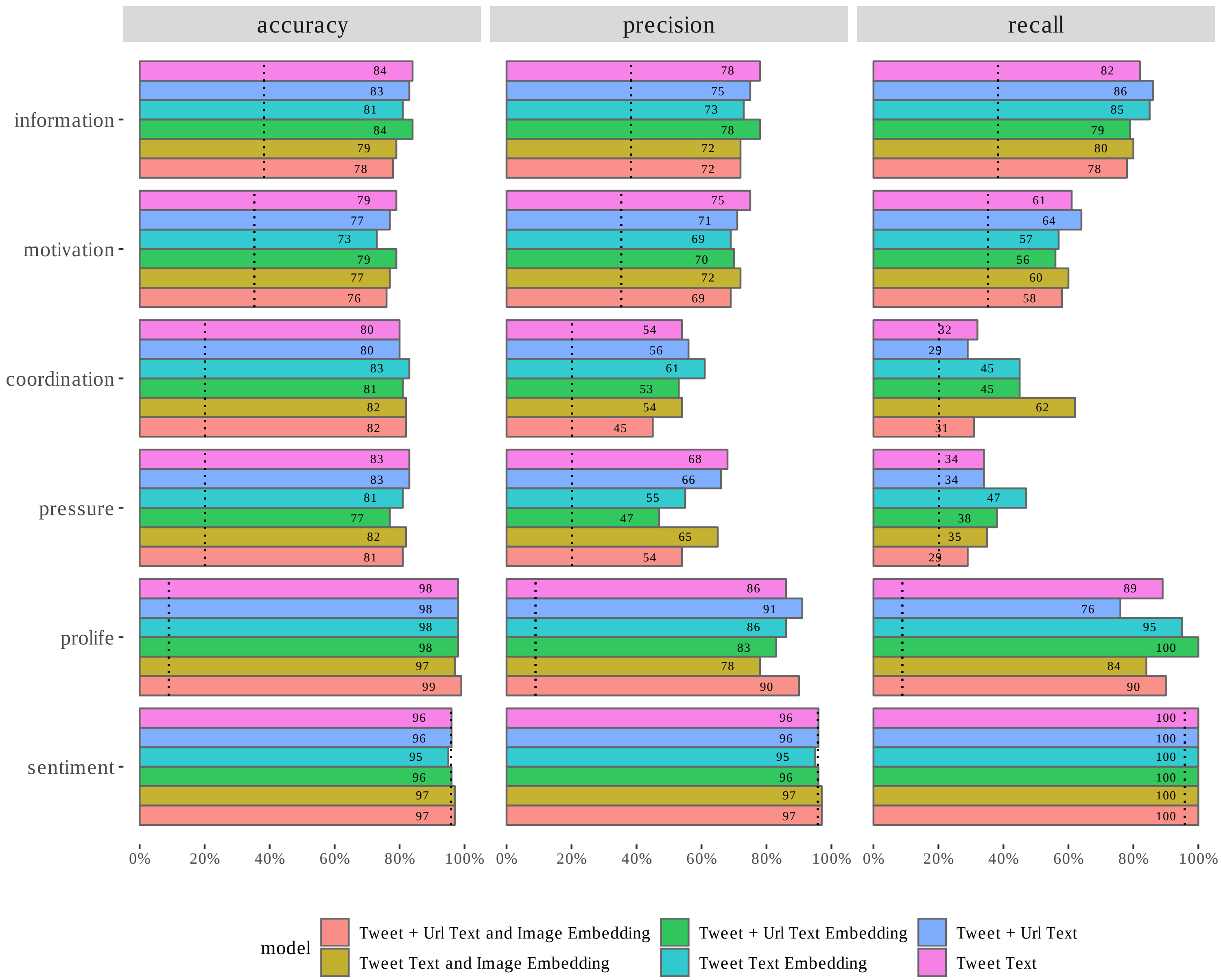
Appendix D Machine Learning Models

Table 6: Description of 6 types of machine learning models trained to predict 6 binary outcomes: the 4 mobilizing mechanism and the two variables of interest (pro-life and sentiment).

Model Label	**Model Type**	**Input Type & Preprocessing**
Tweet Text	Naive Bayes	Document Term Matrix (#Docs $\times$ #Unigrams) • use text from tweet (but not text from shared links) • remove urls (*http*) • remove mentions (@) • remove retweet marks (RT) • all text to lower case • remove all punctuation but exclamation and interrogation signs (!?) • remove stopwords (list from `python nltk` module) • stem all words (*Porter* stemmer from `nltk` module) • same token for all time markers (*timemarker*) • same token for all day markers (*daymarker*) • tokenize and use unigrams to build Document Term Matrix
Tweet + Url Text	Naive Bayes	Document Term Matrix (#Docs $\times$ #Unigrams) • use text from tweet • use metadata text from shared links (Title and Description; if appropriate Twitter cards are present) • apply same preprocessing as in the *Tweet Text* model

Tweet Text Embedding	Logistic Regression	Document Embedding Matrix (#Doc × 100) • use text from tweet • same text preprocessing as in previous models • use pre-trained Word2Vec (Mikolov et al. 2013) model to get 100-size embeddings for each word • average word embeddings within document to get 100-size Doc2Vec embeddings • stack Doc2Vec embeddings to create a Document Embedding Matrix
Tweet + Url Text Embedding	Logistic Regression	Document Embedding Matrix (#Doc × 100) • use text from tweet • use metadata text from shared links • apply same preprocessing as in the *Tweet Text Embedding* model
Tweet Text and Image Embedding	Logistic Regression	Document Embedding Matrix (#Doc × 612) • use text from tweet • apply same text preprocessing as in the *Tweet Text Embedding* model • if tweet has an image, pass it through ResNet18 and pull output from second-to-last layer to get a 512-size image embedding • otherwise, build a 512-size image embedding with 0s • append image embedding to the 100-size text embedding, creating 612-size document embeddings • stack document embeddings to create a Document Embedding Matrix
Tweet + Url Text and Image Embedding	Logistic Regression	Document Embedding Matrix (#Doc × 612) • use text from tweet • use metadata text from shared links • apply same text preprocessing as in the *Tweet Text and Image Embedding* model

Figure 12: Accuracy for an ensemble of 6 machine learning models predicting 6 different outcomes.



In Table 6 we provide a description of the six models trained to predict the 6 tweet-level binary outcomes of the study (the 4 mobilizing mechanisms, plus the anti-march and prolife variables). In Figure 12 we report the accuracy of the models. The dotted vertical lines indicate the percentage of the 200 training tweets that have been found to contain the mechanism. We observe that in most cases overall accuracy is around or above 80%. In terms of precision and recall, for each of the mechanisms there is always a model that yields a precision and a recall above 60% (with the only exception of the recall for the pressure mechanism). This means that the accuracy is satisfactory given that some of the mechanisms are not very frequent (for example, only around 20% of the tweets have been labeled as containing the pressure and coordination mechanism).