Pontifícia Universidade Católica
do Rio de Janeiro

Felipe de Almeida Alvarenga Pereira

The Mechanisms Behind the Effect of Social Media on Protesting Behavior: Evidence from Brazil

Dissertação de Mestrado

Dissertation presented to the Programa de Pós–graduação em Economia of the Departamento de Economia do Centro de Ciências Sociais da PUC-Rio. as a partial fulfillment of the requirements for the degree of Mestre em Economia

Advisor: Prof. Pedro Carvalho Loureiro de Souza

Rio de Janeiro
April 2017

Pontifícia Universidade Católica
do Rio de Janeiro

Felipe de Almeida Alvarenga Pereira

The Mechanisms Behind the Effect of Social Media on Protesting Behavior: Evidence from Brazil

Dissertation presented to the Programa de Pós–graduação em Economia of the Departamento de Economia do Centro de Ciências Sociais da PUC-Rio. as a partial fulfillment of the requirements for the degree of Mestre em Economia

Prof. Pedro Carvalho Loureiro de Souza
Advisor
Departamento de Economia – PUC-Rio

Prof. Gabriel Lopes de Ulyssea
Departamento de Economia – PUC-Rio

Prof. Rudi Rocha de Castro
Universidade Federal do Rio de Janeiro – UFRJ

Prof. Francisco Junqueira Moreira da Costa
Escola Brasileira de Economia e Finanças – FGV/EPGE

Prof. Monica Herz
Coordinator of the Centro de Ciências Sociais – PUC-Rio

Rio de Janeiro, April 7th, 2017

Felipe de Almeida Alvarenga Pereira

B.A., Economics, Universidade de Brasília (UnB), 2014.

Bibliographic data

Pereira, Felipe de Almeida Alvarenga

The Mechanisms Behind the Effect of Social Media on Protesting Behavior: Evidence from Brazil / Felipe de Almeida Alvarenga Pereira; advisor: Pedro Carvalho Loureiro de Souza. – Rio de Janeiro: PUC-Rio, Departamento de Economia, 2017.
(Em Inglês)
v., 60 f: il. color. ; 30 cm

1. Dissertação (mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Economia.

Inclui bibliografia

1. Economia – Teses. 2. Protestos. 3. Mídias sociais. 4. Ação coletiva. 5. Brasil.
I. de Souza, Pedro Carvalho Loureiro. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Economia. III. Título.

CDD: 620.11

Acknowledgments

I would like to thank my parents, Cláudia and Selmo, for their immeasurable support. I am also thankful for all the support, in and out of class, given by my friends and colleagues during these 27 unforgettable months. Financial support from *CAPES* is gratefully acknowledged.

Abstract

Pereira, Felipe de Almeida Alvarenga; de Souza, Pedro Carvalho Loureiro (Advisor). **The Mechanisms Behind the Effect of Social Media on Protesting Behavior: Evidence from Brazil**. Rio de Janeiro, 2017. 60p. Dissertação de Mestrado – Departamento de Economia, Pontifícia Universidade Católica do Rio de Janeiro.

Waves of protests across the world have been linked to the dissemination of social media. In this paper I investigate the effect of social media on protests during the Brazilian protests of June 2013. I exploit the high frequency and short time dimension of the events in order to identify the effect of social media on protests and show that social media activity, through Twitter, positively impacted protesting behavior - attendance and occurrence. I find, for the preferred specifications, that a 10% increase in Twitter activity led to an increase of 6.7% in the number of protestors in the streets and an increase of 3% on the probability of the occurrence of a protest. Furthermore, by analyzing the dynamics of content shared between users, I am able to differentiate between two mechanisms, information diffusion and coordination. Results indicate that more precise coordination was driving the protests through social media, and that information diffusion did not play a role.

Keywords

Protests; Social media; Collective action; Brazil;

Resumo

Pereira, Felipe de Almeida Alvarenga; de Souza, Pedro Carvalho Loureiro. **Mecanismos por trás do efeito de mídias socias sobre manifestações sociais: o caso Brasileiro**. Rio de Janeiro, 2017. 60p. Dissertação de Mestrado – Departamento de Economia, Pontifícia Universidade Católica do Rio de Janeiro.

Ondas de protestos pelo mundo têm sido conectadas à disseminação do uso de mídias sociais. Neste artigo, eu investigo os efeitos de mídias sociais sobre os protestos brasileiros de junho de 2013. Para tal, eu exploro o fato de que os eventos aconteceram em alta frequência e em um curto espaço de tempo para identificar o efeito de mídias sociais sobre protestos e mostro que atividade em mídias sociais, através do Twitter, teve impacto positivo sobre protestos, tanto na margem intensiva quanto na margem extensiva. Encontro, para as especificações preferidas, que um aumento de 10% de atividade no Twitter aumenta em 6.7% o número de manifestantes nas ruas e em 3% a probabilidade de um protesto ocorrer. Além disso, ao analisar o conteúdo compartilhado, e a dinâmica das trocas de informação, consigo identificar dois mecanismos por tras desse efeito: difusão de informação e coordenação. Os resultados indicam que mídias sociais afetaram os protestos ao possibilitar uma melhor coordenação entre os indivíduos, e que difusão de informação não foi relevante.

Palavras-chave

Protestos; Mídias sociais; Ação coletiva; Brasil;

Table of Contents

List of Figures

Lista of Tables

1
Introduction

Recent spread of social media connectedness has fed a significant amount of discussion and research, both in the media and the academia, about its capability of triggering demonstrations and increasing protest participation (1, 2). Horizontal flow of information provided by social media allows people to update their information constantly, through a direct exchange of information with thousands of connections. This information exchange comes at a very low cost, real time, and in a (somewhat) secure environment, where individuals do not feel threatened to express political grievances and affinities. These characteristics of social media potentially enhances coordination between a large number of individuals in many dimensions, including political activism and protesting.

During the events of the Arab Spring, protestors were no longer relying on traditional methods to promote and participate in demonstrations (newspapers, unions, television), but were using social media instead (3). (4)'s tweet exemplifies what was happening: "*We use Facebook to schedule the protests, Twitter to coordinate, and YouTube to tell the world. #egypt #jan25*". Authorities then realized that their control over the traditional media was no longer enough to contain social distress. They started to shut down internet services (5) in an attempt to block any type of mass coordination and, therefore, increase uncertainty among more isolated individuals. Despite the amount of anecdotal evidences, does social media increase protest participation?

In this work I present evidence that social media was relevant in triggering and increasing protests across Brazil in June 2013. I build on a unique data set of social media activity, that is a high frequency short panel of the universe of tweets that contain hashtags related to the protests of June 2013. I use Twitter as a proxy for social media due to its importance among Brazilian social media users. Despite having less than 3% of the world's population, Brazil was the second biggest market for Twitter, with 8% of the total number of users in 2013 (statistics range from 33 to 40 million monthly active users). [1]

[1]Alongside, Twitter is broadly used in the social sciences literature when associating social media with different social outcomes. For instance, (3) concludes that Twitter was the medium of choice for activists in the Iranian protests of June 2009, (6) estimate ideological preferences using connections between users and (7) measure labor market flows through

To deal with the endogeneity in the relationship between social media and protest behavior – attendance or occurrence – I make use of cross section and time fixed effects. The identifying assumption is that, conditioned on city and day fixed effects, social media activity is orthogonal to unobserved determinants of protest participation. This, coped with the high frequency nature of my data, allows me to overcome a problem of omitted variables at the municipality level, such as municipalities with varying degree of political participation or inclination towards the usage of social medias, or common daily events, such as a presidential announcement. This identification strategy, however, fails to identify a causal effect if reverse causality is present in the relationship between social media activity and protesting.

I study two potential mechanisms behind the effect of social media on protests. First, the informational channel (8). Here, connected individuals receive an influx of relevant information regarding the protests and then decide, upon this information, whether to protest. Such information can be of two types, either knowledge about the existence of a protest, or relevant information against the main target of the demonstrations, which in this case was the government, that could trigger responses from individuals. I use timing and content of Twitter data to identify whether information diffusion, and what kind, was driving the protests.

The second mechanism is the coordination channel (9, 10), or the reduction of collective action costs. Here, connected individuals infer, with more precision, information on place, time, expected level of violence and number of expected protestors who will be attending. With this reduced uncertainty, coordination would bring more people to the streets. However, reduced uncertainty through coordination could come through different deeper channels. Knowing whether protestors coordinated before or during the demonstrations helps understanding if social media induced more impulsive protestors to the streets, or fostered discontent for a period of time before taking the streets, for example. With activity data on the timing of shared information I can identify three types of coordination: (i) a “long term” planning, where users coordinate ahead within an interval of days; (ii) pre-protest, same day coordination, where users share information on time and place hours before the protest; and (iii) live logistical coordination, where users inform on protesting obstacles such as police presence, level of violence and blockades during the protest.

Overall, the results I obtain reveal a positive impact of social media activity on protests. For my preferred specifications the effect of a 10% increase in Twitter activity leads to 6.7% increase in the average number of protestors.

labor market related Twitter posts.

The same occurs for the extensive margin, where a 10% increase in Twitter activity leads to a 3% increase in the probability of a protest to happen. Moreover, results indicate that the Brazilian demonstrations were affected by social media through a pre-protest, same day coordination, and that the information diffusion mechanism did not play a role in enhancing protest attendance through social media.

The Brazilian setting is an interesting event study for a few reasons. First, Brazil was considered a stable democracy at the time, which is not the usual setting where the relationship between social media and protests is studied. Most studies focus on countries under authoritarian rule (11, 12) or that have a history of media censorship and/or manipulation (5, 13). In such environments, the role of social media is boosted, since individuals turn to it as the safest means to consume and generate information (14). In a setting without any clear restrictions on the role of traditional media on reporting information about protests and criticisms against the government, social media should have a different role than in censored environments. Competition from traditional media should lower the importance of social media on information diffusion, for example, diminishing its impact on the protest attendance through an informational channel. Identifying an important role of social media on protesting behavior in a country without clear constraints on media, where it suffers direct competition from other types of media, provides evidence of a unique feature of social media in affecting social demonstrations.

Second, Brazilian protests had both time and geographic variability, which provides sufficient observations for an econometric analyzes, one of the considerable difficulties in studying this relationship.[2]. Moreover, it also provides a framework that allows for a fixed effects identification strategy.

There has been extensive work on the coincidence of the rise in social media diffusion and protest behavior. Positive correlations between social media penetration (number of users), or activity, and protest behavior have been reported and interpreted (17, 18, 16, 12). However, it is important to stress that there is no consensus as for the direction of the effect. Initially one would think that better information and communication technologies (ICTs) should tilt the balance toward protestors, by, for instance, raising the costs of media manipulation by the governments. Yet, evidences have been provided for the opposite. (19) shows that West Germany television broadcasts

[2]Most of the work done in correlations between social media and protests analyzed events that happened within a small time frame, or had long pauses between events, rendering 100 or 200 days of observations (15, 16) or even less than 60 days of observations (17). Spatial variation is hard to find since most events happened in small countries or were concentrated in a few cities. Aggregating different countries is often the answer (12)

increased East Germans' support for East Germany's regime, instead of creating or boosting an anti-government awareness. Alongside, (20) find that broader cellphone coverage undermined Iraqi's anti government movements' coordination and communication, instead of improving them, by allowing the government to predict and dismantle dissident organizations.

Recent work by (13) finds that social media penetration did have a positive impact in the Russian protests of 2011-2012. They use the place of birth of students who attended college together with the founder of *VKontakte* (a Russian social media) as an exogenous variation to identify the effect of this social media penetration, across cities, in the Russian protests. Their results indicate that the mechanism behind the impact of social media penetration was the reduction of collective action costs. Roughly, a 10% increase in the number of *VKontakte* users in a city would increase protest participation by approximately 20%, on average.

This paper also relates to a rising literature on the relationship between the diffusion and development of ICTs and social outcomes. Interesting findings provide evidence on the effect of radio transmission on violence (21), internet diffusion on political turnout (23, 22), diffusion of mobile phones on protests (24) and television diffusion on the ruling party vote share (25).

Overall, relative to the existing literature, my contributions are twofold. First, I identify a positive effect of social media activity on protest behavior, on a country with a stable democracy and no record of systematic media censorship. Second, I am able to distinguish between more precise mechanisms through which social media affects protesting behavior, finding that the timing of social media activity is important to explain the Brazilian protests.

The rest of the article is structured as follows. Chapter 2 characterizes the event study. Chapter 3 describes the data. Chapter 4 discusses the empirical strategy. Chapter 5 presents the results. Chapter 6 discusses the mechanisms. Chapter 7 concludes.

2
Context

In June 2nd, both the City and the State of São Paulo decided to increase public transportation fares by 20 cents. Governments defended the increase by saying that there had been no adjustments in the tickets for over two years, while costs and inflation were running high.

Upon this decision, *Movimento Passe Livre* (Free Fares Movement) called for protests against the raise in transportation fares, and scheduled it for June 6th (26). Two thousand people attended the calling and went to the center of São Paulo to block the streets and call for the cancellation of the increase. However, protestors also vandalized metro stations and bus stops. Police tried to contain the demonstration, with a strong show of force and violence, which ignited protestors to schedule more protests (27).[1]

In June 13th the movement gathered for another protest, now both in São Paulo and Rio de Janeiro. This event witnessed, again, violence from both sides, however, police violence was more pronounced, disposing of rubber bullets and tear gas bombs. Of the estimated five thousand people on the streets in São Paulo, police detained 160 protestors and 55 people were injured, 7 of them journalists. In Rio de Janeiro, 2,000 people took the streets and, at the end, police displayed the same tactics as in São Paulo (28). The excessive police violence wasn't taken smoothly by the country.

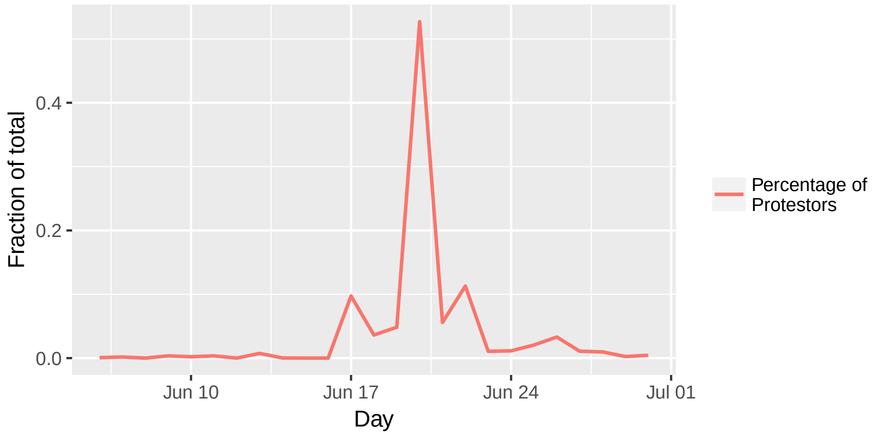
Drawing upon the violence displayed by the police, the protests, that claimed for reduced transportation costs, became national demonstrations against how government treated its citizens (35). Specially after news from government expenditures on new stadiums for the FIFA World Cup, to be held the next year, started to flow the internet. Protestors started to ask the government for "FIFA standards" on public goods and services. These claims followed what was called "[...] an immense and obscene expenditure of public resources" (29) and the government justification that it had to meet the FIFA standards of quality for the World Cup.

The movement itself quickly escalated and disseminated discontent throughout the country. In a span of 25 days it reached over 340 different muni-

[1] Media, government and people criticized methods used by the protestors and called the demonstrations vandalism (27).

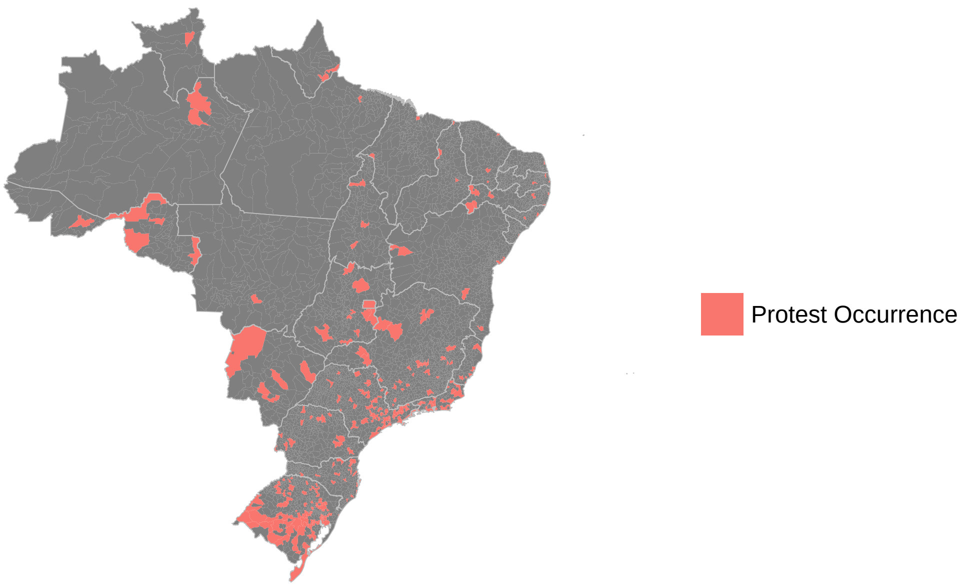
cipalities and brought over 2.8 million people to the streets. Figure 2.1 shows the evolution of the number of people on the streets through time and figure 2.2 shows the geographic penetration of the protests.

Figure 2.1: Daily percentage of total number of protestors through time



Notes: Each observation refers to the amount of the specific day divided by the sum of the amount of all days. Source: *G1* news website.

Figure 2.2: Municipalities that experienced at least one protest in June 2013



Source: *G1* news website.

Municipalities that experienced protests were quite different from the rest of the country, despite having a lot of heterogeneity among them. Table 2.1 shows descriptive statistics for the municipalities that had, and had not, experienced at least one protest event during the 25 days. The only variable that is not rejected in the test on the difference of means is the proportion of young people in the municipalities. Cities that experienced protests were richer, better educated and better developed in terms of infrastructure.

Table 2.1: Descriptive statistics – Difference of means

	Cities with protests Mean	Cities without protests Mean	Difference of means p-value
Population 2013	271,970.2	27,611.65	0
	(810,385.89)	(48,380.24)	
GDP 2013 per capita (R$)	30,253.31	18,332.89	0
	(22,210.3)	(21,824.82)	
% with HS diploma or higher	35.9%	20.75%	0
	(9.58)	(7.54)	
% Men	49.04%	50.59%	0
	(1.1)	(1.55)	
% living in urban areas	89.14%	62.17%	0
	(13.07)	(21.49)	
% HH with Internet Connection	31.9%	13.66%	0
	(11.26)	(10.26)	
% HH with cellphones	86.54%	72.06%	0
	(5.5)	(16.05)	
% HH with electricity	97.63%	94.85%	0
	(2.5)	(8.6)	
% HH with sewer coverage	53.64%	27.67%	0
	(29.14)	(30.58)	
% HH with clean water	94.51%	82.88%	0
	(5.93)	(18.75)	
% people between 10-24 years old	6.08%	6.06%	0.865
	(1.59)	(2.11)	
% people between 25-39 years old	30.32%	29.51%	0.001
	(4.4)	(4.83)	
Distance to São Paulo	742.39 km	1190.23 km	0
	(605.52 km)	(753.73 km)	
Number of cities	341	5224	

Notes: Standard deviations in parenthesis. p-value column refers to the p-value of the t-test on the difference of means, for the two groups of cities, first and second columns, under a 5% threshold.
All variables come from IBGE 2010 Census except GDP per capita and Population, which come from IBGE annual track. GDP per capita is measured in annual Reais. Distance to São Paulo was measured in a straight line from the city centroid to the remaining municipalities' centroids. HH corresponds to households.

Looking specifically to the cities that experienced protests, (30) shows, through controlled regressions, that there was heterogeneity in a few dimensions among these cities. They show, for example, that the municipalities that had more unsatisfied people regarding health and education public services, and cities that had higher levels of corruption perception had, on average, more protests.[2]

Aside from these characteristics, the June protests were unique in two important dimensions. First, they were short and happened in a high frequency, with protests happening every day in a span of approximately 25 days. It can be seen from figure 2.1 that at the end of June the movement was quite smaller than at its peak. After June 30th only scattered demonstrations happened, without any resemblance to June's protests.

[2]See also (30) for descriptive statistics on the difference between cities that did not have protests and those which had.

Second, the movement lacked a concrete leadership coordinating the protests. Both of the previous nationwide movements, 1983-84 and 1992, had either political parties or workers' unions drawing on the unfolding of the events (26). June 2013 had no leading figure, political party or institution that helped trigger the protests. After halfway through the demonstrations a few online groups received more attention, but, still, without any of the conventional leaderships being decisive in the protest development.[3]

Anecdotal evidences suggest that social media filled the role of organizing the demonstrations. Hashtags from these demonstrations always took a high rank in the Twitter trending topics for Brazil, as well as internationally, during the events. Information on event planning, the expected level of violence, and even live coverage, was being shared throughout the network (32), as well as propaganda against the government (26).

However, for social media to have taken the role of triggering and increasing protests, first it needed to be disseminated across the population. Indeed, the Brazilian network of social media users, in 2013, was quite relevant. In terms of size, Brazil was Twitter's second largest market, with between 33 and 40 million active users. *Latinobarómetro* survey shows that 48.2% of Brazilians used some kind of social media in 2013 (up from 30% in 2010).

This large network of users coped with the fact that Twitter was broadly used for real time news consumption (33), and news production (34), makes a suitable environment for an analysis of the effects of social media in protesting behavior.

[3]See (31) for a more comprehensive analyzes of the Brazilian protests of June 2013 and its effects on political outcomes.

3
Data

Twitter data was obtained through the GNIP application program interface (API), Twitter's branch that stores all data generated by its users. The collected dataset was the universe of tweets that contained at least one hashtag from a set of 157 hashtags related to the Brazilian protests. Hashtags were used to identify tweets talking about the protests due to their capacity to uniquely reference a specific subject and to avoid problems such as miss-correspondence, that could occur if tweets were chosen through text mining in their content. This set of chosen hashtags were hand selected and attempt to cover the broadest possible range of protest tweets during the events of June 2013. Table A.1 in the appendix lists all hashtags used to identify protest tweets.

To use this data to analyze the effects of social media on protesting behavior I need to identify the locations where these tweets were posted. Twitter data generated by its users allows for two types of location identification. If a user tweets from a mobile device, and allows Twitter to access its GPS information, coordinates from each tweet are shared with the platform, which enables me to pinpoint the user location at the moment he shared information. Geo-coded coordinates can come as a set of 4 points, generating a polygon around a specific area - which I call area referenced -, or a single point in space - which I call point referenced -, depending on the quality of the GPS device. A second possibility is to exploit the user self-reported location. Since providing any location on Twitter is not mandatory, only 63% of the tweets have a declared, either self-reported or geo-coded coordinates (or both). Table 3.1 shows how much of the data set has some type of user location.

Table 3.1: Tweets in Numbers

From 06/06 to 06/30		Total	As fraction
All Tweets		2,323,658	1
With Location	Point referenced	44,821	0.019
	Area referenced	21,999	0.009
	Self reported	1,401,199	0.60

Although geo-coded coordinates are better than self-reported locations in terms of precision (there are no rules for users to report their locations, allowing them to report imaginary cities or jokes as their location), only 3% of the data provide such coordinates. Therefore, I aggregate tweets from both geo-coded and self-reported locations to increase the number of observations that can be located. [1] I do not consider tweets posted outside of Brazil, which reduces the number of *Point* and *Area* referenced tweets from 66,820 to 63,489. Out of the 1,401,199 tweets that can only be identified with a self-reported location, only 716,587 of them matched any of the existing Brazilian cities in 2013. I show in the appendix robustness checks using only geo-coded tweets and show that all results stay stable. The matching procedure of the algorithm is described in the appendix subsection A.2.

To measure the efficiency of the matching algorithm I look at tweets with both self-reported locations and geo-coordinates, which amounts to approximately 50k tweets. I then compare their geo-locations to the matches produced under the matching algorithm. This approach reveals that 85.3% of the geo-locations match the user self reported location. The tweets that did not match the city, matched the state in 98% of the cases, all of them in cities that were at a driving distance from the geo-located tweet. These unmatched tweets probably reveal people that live and work at different places. Therefore, with these numbers, it is reasonable to assume that the algorithm was successful.

Protest data was hand-collected through news mining. Most of the data came from the compilation done by the news website *G1*, which provided information on the estimated number of protestors for each Brazilian municipality and each day of protests.[2]

Table 3.2 summarizes the information that will be carried out for the empirical analysis, and table 3.3 shows the difference of means of specific characteristics between cities. Panel A shows the difference of Twitter activity between cities that had protests and those that had not. The former experienced over 50 times more tweets than the latter. Panel B shows descriptive statistics between cities that had and had not Twitter activity. Cities are different in all dimensions, except for the percentage of people between 10 and 24 years old.

After aggregating tweets by day and city, alongside protest data, I create a panel on the number of protestors and protest-related Twitter activity in a given city-day. The final data set will then consist of a high frequency short

[1]Using self-reported locations is usual in the context of social media analysis (13).

[2]G1 is the online platform of the biggest media group in Brazil, *Globo*. The protest compilation website can be accessed through: *http://g1.globo.com/brasil/protestos-2013/infografico/platb/*. Most of their data were provided by state police statistics. This can be an issue if police misreported data intentionally for image purposes. If this measurement error occurred, it was, probably, more pronounced in small cities, where there is no presence of different statistical institutions to run independent statistical assessments.

Table 3.2: Data set for tweets with identifiable locations

	Tweets with location	Protestors
Total	780,076	2,836,050
Cities	3,672	341
Days	25	25

Notes: Tweets with location aggregates both self-reported data and GPS data (Area + Point referenced tweets).

Table 3.3: Descriptive statistics – Difference of means

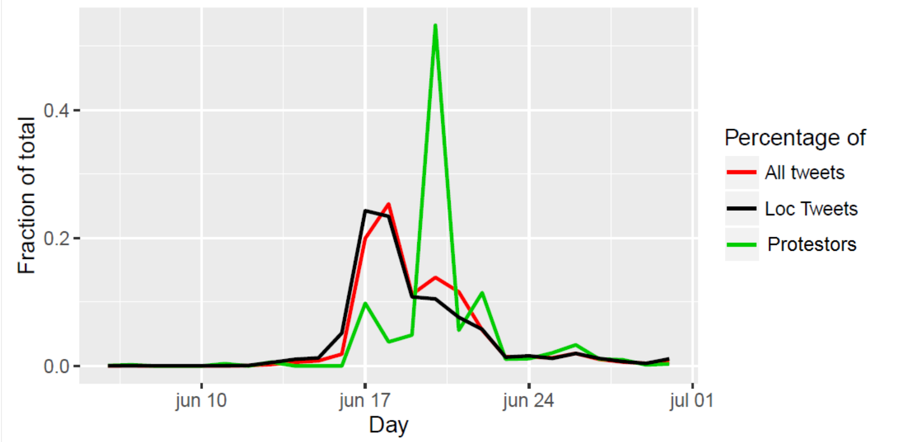
Panel A	Cities with protests Mean	Cities without protests Mean	Difference of means p-value
Protests tweets	1,936.54	35.87	0.001
	(10,647.21)	(128.79)	
Protest tweets per capita	0.003	0.002	0.0003
	(0.003)	(0.018)	
Number of cities	341	5224	
Panel B	**Cities with tweets Mean**	**Cities without tweets Mean**	**Difference of means p-value**
Population 2013	50,347.42	8,615.86	0
	(260,868.9)	(7,893.73)	
GDP 2013 per capita (R$)	19,445.19	13,406.37	0
	(22,146.09)	(14,561.54)	
% with HS diploma or higher	23.93%	17.29%	0
	(8.87)	(5.54)	
% Men	50.18%	51.11%	0
	(1.42)	(1.66)	
% living in urban areas	69.78%	52.31%	0
	(20.84)	(19.61)	
% HH with Internet Connection	17.97%	8.54%	0
	(11.55)	(7.27)	
% HH with cellphones	76.66%	65.66%	0
	(13.61)	(17.73)	
% HH with electricity	96.2%	92.71%	0
	(6.5)	(10.82)	
% HH with sewer coverage	35.7%	16.67%	0
	(31.92)	(25.12)	
% HH with clean water	86.84%	77.24%	0
	(15.68)	(21.53)	
% people between 10-24 years old	6.06%	6.06%	0.904
	(1.94)	(2.33)	
% people between 25-39 years old	29.69%	29.3%	0.006
	(4.63)	(5.14)	
Distance to São Paulo	1,062.57 km	1,358.78 km	0
	(745.8 km)	(728.68 km)	
Number of cities	3668	1897	

Notes: Standard deviations in parenthesis. p-value column refers to the p-value of the t-test on the difference of means, between the two groups of cities, first and second columns, under a 5% threshold.
Whole set of tweets identified through hashtags amount to 2.3 million, however, not all of them can be matched to an existing location. All variables come from IBGE 2010 Census except GDP per capita and Population, which come from IBGE annual track. GDP per capita is measured in annual Reais. Distance to São Paulo was measured in a straight line from the city centroid to the remaining municipalities' centroids. HH corresponds to households.

panel of 25 days and 3,672 cities. Figure 3.1 shows the evolution of protest-related tweets and number of protestors, as percentage of the total during the

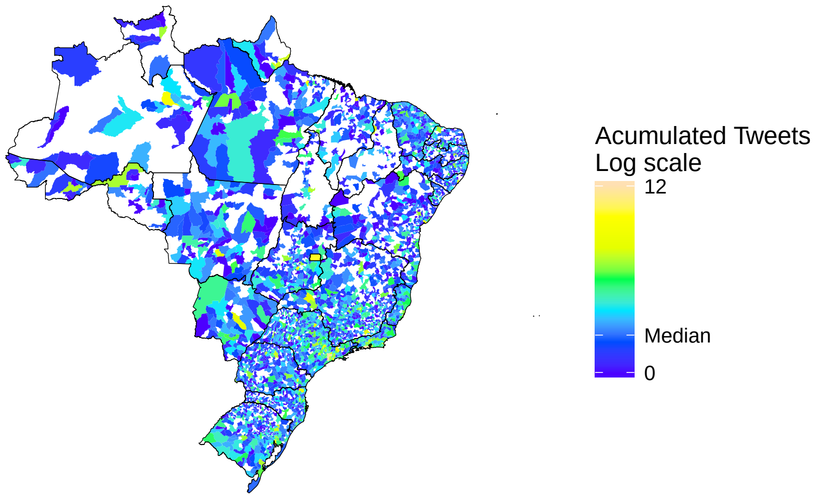
25 days of protests, and figure 3.2 shows the intensity of Twitter activity across municipalities, for the entire period.

Figure 3.1: Daily percentage of total tweets and protests through time



Notes: Each observation refers to the amount of the specific day divided by the sum of the amount of all days. *E.g.*, for the black line, June 17th shows the amount of tweets with location on the 17th divided by the total number of tweets with locations for the entire period.

Figure 3.2: Intensity of Twitter activity for the entire period

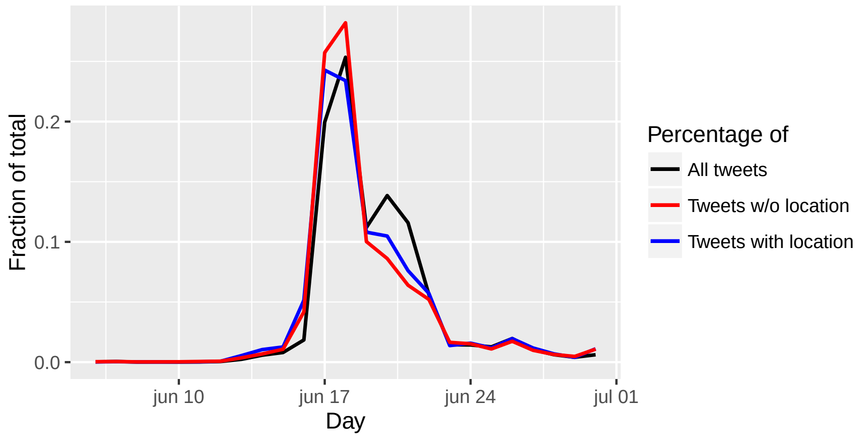


Notes: Data plotted is the log of the sum of tweets over time, for each municipality.

Tweets with identifiable locations represent roughly a third of the universe of protest-related tweets. Figure 3.3 shows the time series of the events for tweets with and without locations, as well as the total amount. The data plotted is the number of tweets in a day divided by the total number of tweets for the entire period. The graph reveals that both groups, tweets with or without location, follow the same trend, with minor deviations in June 17th and 18th. From this point of view, figure 3.3 suggests that the selection of *lo-*

cationable tweets is not affecting the behavior of the sample, when comparing the time trends between the two different groups.

Figure 3.3: Daily percentage of total tweets through time - location vs no location



Notes: Each observation refers to the amount of the specific day divided by the sum of the amount of all days. *E.g.*, for the blue line, June 17th shows the amount of tweets with location on the 17th divided by the total number of tweets with locations for the entire period .

Given that I cannot locate the other 2/3 of the data set, I cannot do a similar analysis for the geographic dispersion. Therefore, it is impossible to assess whether the selection of *locationable* tweets affected the distribution of tweets across municipalities. However, I can analyze differences in users' characteristics between the data I can locate and the one I can't, in order to indirectly infer whether selection is affecting the geographic distribution of my sample.

Table 3.4 informs on how different these users are in four dimensions. Users are statistically equal in their average number of followers and lifetime posts. However, they are not in terms of average number of accounts followed and protest-related posts.

Even though these numbers do not provide direct information on the geographic dispersion of these users, the fact that the average lifetime Twitter activity is statistically equal between users with and without a declared location is by itself interesting. If Twitter activity is correlated to the users' location, then the averages being equal between these two types of users indicates some correlation for the geographic dispersion between them.

The fact that I can only identify locations for a third of all protest-related tweets is important for the subsequent procedures in this article. Since my object of analysis is the amount of tweets in a city for a given day, having less tweets than the actual amount generates higher point estimates then the

Table 3.4: Descriptive user statistics by location status - differences of means

	Tweets with location	Non-Loc tweets	S.d.	p-value
Following	433.54	461.20	7.241	0.001
Followers	1,270.56	1,177.25	90.058	0.30
Protest posts	3.94	3.50	0.036	0.001
Accum. lifetime posts	9,838.36	9,809.07	43.380	0.50
Tweets†	780,076	1,464,543		
Users	197,733	417,630		

Notes: † Tweets outside Brazil, either in self-reported or GPS locations, have been dropped. p-value column refers to the test on the difference of means between the first two columns. The values refer to a t-test under 5% confidence.

actual effect. However, knowing the extent of the gap between the universe and my sample of tweets helps on inferring the magnitude of this bias.

4
Empirical Strategy

In this work I estimate whether social media activity, through Twitter, positively affected the June protests in Brazil. In order to estimate a causal effect of social media in protests, I would need social media to be orthogonal to any uncontrolled and unobserved determinants of protests. An equivalent experimental setting would be to randomly deprive a set of municipalities of social media access, by blocking their IP addresses to connect to Twitter, for example. Since this is not feasible, I make use of the special nature of my data set to identify the effect of social media.

Given that the protests started and ended in less than 30 days, and both protest activity and social media activity were happening in a daily basis, the resulting data is a high frequency panel. With that, the employment of time and cross section fixed effects controls for, basically, all unobserved city and time characteristics that could bias the estimation of the effect of social media on protests. Infrastructure, public sector penetration, income and educational differences across municipalities would be controlled for, as well as common daily events such as national announcements by the president on television and the national behavior of traditional media. The short time dimension of the event helps on keeping constant variables such as the municipalities' propensity to protest, which I assume that do not vary in such a small time frame.

The regressions I will estimate will then follow the standard fixed effects panel setting, assuming the following identification hypothesis: conditional on city and day fixed effects, social media is orthogonal to unobserved determinants of protest participation.[1] The regressions I estimate are the following:

$$Protestors_{ist} = \beta Tweets_{ist} + FE_i + FE_t + FE_s * FE_t + u_{ist} \tag{4-1}$$

[1]Still, the whole set of fixed effects do not control for a potential reverse causality. I address this potential problem in the appendix, where I propose an instrumental variables approach.

$$Protestors_{ist} = \beta_1 Tweets_{ist} + \beta_2 Tweets_{ist-1} + \delta Protestors_{ist-1} + \\ + FE_i + FE_t + FE_s * FE_t + u_{ist} \quad (4\text{-}2)$$

where i refers to the city, s to the state and t to the specific day. $Tweets$ is the variable that accounts for the amount of tweets and $Protestors$ accounts for either the number of protestors or an indicator for the existence of a protest. FE stands for fixed effects and u_{ist} is the regression error.[2]

[2]I show in the appendix section A.5 count data regressions using both poisson and negative binomial panel estimations.

5
Results

5.1
Protest behavior

The first empirical exercise that I present, which I call protest behavior, studies the entire data set, analyzing all observations together. In table 5.1 all cities that had either 1 day of protests or 1 day with tweets entered the data set. Panel A shows the level-level specifications and Panel B the log-log specifications.

First two columns of both panels shows the results of regression 4-1, whereas the last two columns show results for the addition of one lag of the dependent variable and one lag of Twitter activity, according to regression 4-2. Odd columns shows coefficients for a regression without any controls and even columns shows the fixed effects specifications.

From table 5.1 we can see that social media activity is significant, both economically and statistically, at all specifications. Joint significance tests are significant below the 1% threshold.

Interpretation of Panel A estimates are straightforward. In column 4, one more tweet increases the number of protestors by approximately 2.4 people, on average. In percentage terms, 10% increase in contemporary tweets for the whole period (78k tweets) would lead to an increase of 6.7% in the number of protestors (186k people). For the log-log specification in Panel B, 10% increase in contemporary tweets would lead to 1.37% increase in the number of protestors (40k people).

Interesting to point out that, for the lags specifications, contemporary Twitter activity dominates past activity in terms of magnitude, both with or without fixed effects. This finding could be related to the specific use people had for Twitter during the June protests, differentiating between information consumption or logistical coordination, for example. I extend on this in chapter 6.

5.2

Table 5.1: Protest Behavior and Twitter activity

	Protestors$_t$			
	(1)	(2)	(3)	(4)
Panel A: Level-Level Regressions				
Tweets$_t$	2.404***	2.305***	2.473***	2.366***
	(0.607)	(0.659)	(0.870)	(0.892)
Tweets$_{t-1}$			0.065	0.065
			(0.390)	(0.424)
Protestors$_{t-1}$			−0.089***	−0.124***
			(0.033)	(0.027)
F stat	29754.28	8.3	9867.93	8.52
p-value	(0.00)	(0.00)	(0.00)	(0.00)
R^2	0.246	0.291	0.252	0.304
Adjusted R^2	0.246	0.256	0.252	0.268
Panel B: Log-Log Regressions				
log(Tweets$_t$)	0.179***	0.165***	0.130***	0.137***
	(0.014)	(0.012)	(0.010)	(0.010)
log(Tweets$_{t-1}$)			0.055***	0.054***
			(0.006)	(0.008)
log(Protestors$_{t-1}$)			0.045**	−0.056***
			(0.023)	(0.018)
F stat	8756.32	4.89	2994.81	4.87
p-value	(0.00)	(0.00)	(0.00)	(0.00)
R^2	0.087	0.195	0.093	0.200
Adjusted R^2	0.087	0.155	0.093	0.159
City FE	No	Yes	No	Yes
Time FE	No	Yes	No	Yes
State*Time FE	No	Yes	No	Yes
Observations	91,350	91,350	87,696	87,696

Notes: Clustered standard errors, adjusted within micro-regions, in parenthesis. All logarithms were calculated adding 1 inside. *p<0.1; **p<0.05; ***p<0.01

Extensive margin

Now I turn to investigate whether social media activity increased the probability of a protest to happen. Here I present the results when the dependent variable is binary, assuming the value of one when there is a protest in a given city-day. Specifications are of the form of a linear probability model. I present logit and fixed effects logit estimations in the appendix section A.6 for robustness checks.

Table 5.2 presents the level-level and level-log specifications of regression 4-2. Odd columns shows the effects without any fixed effects, whereas even columns shows the results with the full set of time and cross sectional fixed effects.

Table 5.2: Linear Probability Model

	Protest$_t$			
	Level-level		Level-log	
	(1)	(2)	(3)	(4)
Tweets$_t$	0.00004***	0.00003**	0.017***	0.017***
	(0.00001)	(0.00001)	(0.001)	(0.001)
Tweets$_{t-1}$	0.00001***	-0.00001	0.008***	0.007***
	(0.00001)	(0.00001)	(0.001)	(0.001)
Protest$_{t-1}$	-0.00000	-0.00001^{***}	0.007**	-0.007^{***}
	(0.00001)	(0.00001)	(0.003)	(0.002)
City FE	No	Yes	No	Yes
Time FE	No	Yes	No	Yes
State*Time FE	No	Yes	No	Yes
F stat	863.15	3.84	2503.36	4.29
p-value	(0.00)	(0.00)	(0.00)	(0.00)
R^2	0.029	0.165	0.079	0.180
Adjusted R^2	0.029	0.122	0.079	0.138
Observations	87,696	87,696	87,696	87,696

Clustered standard errors, adjusted within micro-regions, in parenthesis. All logarithms were calculated adding 1 inside. log(Protestors$_{t-1}$) is the log of the amount of protestors in last period. *p<0.1; **p<0.05; ***p<0.01.

Although the estimated coefficients are low, one has to look at the effect at the average levels of activity to understand the impact of social media on the incidence of protests. A 10% increase in the contemporary number of tweets (78k people) would increase the probability of a protest to happen by approximately 3% for the level-level specification and 1.6% for the level-log.

The effects follow the same pattern found in the protest behavior analysis. Contemporary tweets dominate past activity in terms of magnitude, with past activity being either not significant or less than half the effect of contemporary activity.

5.3
Intensive margin

The next step is to analyze the marginal effect of social media in protest attendance. For that, I subset the data only into municipalities that experienced at least one protest event during the entire period. The resulting set of observations include 341 cities during the 25 days of protest.

Table 5.3 shows the results for regression 4-2, where, again, odd columns do not control for any fixed effects and even columns do.

Table 5.3: Protest Participation and Twitter Activity

	$Protest_t$			
	Level-level		Log-log	
	(1)	(2)	(3)	(4)
$Tweets_t$	2.467***	2.344**	0.332***	0.591***
	(0.874)	(0.929)	(0.025)	(0.039)
$Tweets_{t-1}$	0.062	0.080	0.033	−0.114***
	(0.393)	(0.450)	(0.023)	(0.036)
$Protest_{t-1}$	−0.091***	−0.136***	−0.033	−0.114***
	(0.033)	(0.029)	(0.021)	(0.018)
City FE	No	Yes	No	Yes
Time FE	No	Yes	No	Yes
State*Time FE	No	Yes	No	Yes
F stat	910.52	4.55	400.01	4.03
p-value	(0.00)	(0.00)	(0.00)	(0.00)
R^2	0.250	0.378	0.128	0.350
Adjusted R^2	0.250	0.295	0.128	0.263
Observations	8,184	8,184	8,184	8,184

Notes: Clustered standard errors, adjusted within micro-regions, in parenthesis. All logarithms were calculated adding 1 inside. *p<0.1; **p<0.05; ***p<0.01

Results, again, show that contemporary activity is what is driving protestors through social media. Level-level specifications show, basically, the same coefficients as in table 5.1. However, log-log specifications are now stronger, where a 10% increase in the number of tweets (78K) would generate a 5.9% increase in the number of people in the streets (168k people). This effect is closer to the effect estimated in the level-level specification (6.7%) the one estimated in the protest behavior section.

5.4 Heterogeneous effects

Here I investigate whether the effects of social media were similarly diffused across cities or were more pronounced for specific demographics or ideologies. For that, I run regression 4-2 into different subsets of the data. I separate the main data set used in the estimations for the effect of social media on protest behavior (section 5.1) into six different groups of cities: above and under the median municipalities' GDP, above and under the median share of high school graduates in the population, and cities that had or not the Workers' Party (PT) in the executive office by the time of the protests.[1]

The expectations here are that low income and low educated municipalities' protests were affected less than their counterparts' by social media. The main reason for such expectation is that low income and low educated municipalities should have a more difficult access to internet services and, thus, have less social media penetration. This should happen as a result of the households' incapacity to pay for internet services, or the municipality's low level of development that makes it intractable for internet providers to reach the city's households (*e.g.*, high fixed cost to lay down internet infrastructure).

However, the effect of social media over protests on municipalities that had PT as the ruling party in the executive office is not clear. Broadly, the main target of the demonstrations was the federal government, because of its bad management of public finances and services, as well as pure political discontent against the ruling president, who was affiliated to PT (35). Municipalities ruled by PT mayors could have a population more inclined to forgive government misdoings and defend their mayors on social media, or simply disregard protest related social media activity. On the other hand, those cities could be a more attractive target of discontent, by having ties with the main national grievance, and have more pronounced demonstrations and social media activity.

Table 5.4 shows the results of the effect of social media on protest behavior for each of the six subsets. Panel A shows level-level regressions and Panel B shows the logs specifications. The numbers confirm the expectations for income and education regressions: more educated and richer cities have a high and significant effect of social media on protest behavior, whereas the poorer and less educated have either coefficients statistically zero or over ten times lower than the riches cities.

As for the ideological separation, Panels A and B disagree. For Panel

[1]PT was the party of the ruling president, Dilma Roussef, by the time of the protests. Since the federal government was heavily targeted during the demonstrations, having ties with it could generate different effects of social media over protesting behavior.

A, municipalities ruled by mayors not affiliated to PT had an effect of social media three times higher than cities which were ruled by PT mayors. However, Panel B specifications says the opposite.

Table 5.4: Heterogeneous effects - intensive margin

	Protestors$_t$					
	High educ	Low educ	High income	Low income	PT	Not PT
	Panel A: Level-Level Regressions					
Tweets$_t$	2.363***	0.019*	2.367***	0.001	1.165***	3.863***
	(0.897)	(0.011)	(0.896)	(0.001)	(0.037)	(0.190)
Tweets$_{t-1}$	0.066	−0.003	0.066	−0.0005	0.864***	−0.661***
	(0.427)	(0.004)	(0.427)	(0.001)	(0.100)	(0.108)
Protestors$_{t-1}$	−0.125***	−0.047***	−0.125***	−0.043***	−0.331***	−0.096***
	(0.027)	(0.002)	(0.027)	(0.001)	(0.059)	(0.028)
F stat	7.61	1.16	7.53	0.9	11.5	9.14
p-value	(0.00)	(0.00)	(0.00)	(0.997)	(0.00)	(0.00)
R^2	0.311	0.063	0.308	0.050	0.546	0.324
Adjusted R^2	0.270	0.009	0.267	−0.005	0.498	0.288
	Panel B: Log-Log Regressions					
log(Tweets$_t$)	0.188***	0.010**	0.193***	0.002	0.195***	0.128***
	(0.014)	(0.004)	(0.014)	(0.003)	(0.026)	(0.011)
log(Tweets$_{t-1}$)	0.053***	0.001	0.052***	−0.0005	0.029	0.057***
	(0.010)	(0.002)	(0.010)	(0.001)	(0.018)	(0.008)
log(Protestors$_{t-1}$)	−0.064***	−0.050***	−0.066***	−0.044***	−0.130***	−0.040*
	(0.018)	(0.004)	(0.018)	(0.001)	(0.018)	(0.021)
F stat	4.73	1.15	4.72	0.89	2.99	4.75
p-value	(0.00)	(0.00)	(0.00)	(0.998)	(0.00)	(0.00)
R^2	0.219	0.063	0.218	0.050	0.238	0.199
Adjusted R^2	0.173	0.008	0.172	−0.006	0.158	0.157
City FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
State*Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	43,824	43,872	43,848	43,848	10,896	76,800

Notes: Clustered standard errors, adjusted within micro-regions, in parenthesis. All logarithms were calculated adding 1 inside. *p<0.1; **p<0.05; ***p<0.01.

Table 5.5 shows the same subset regressions, but looking at the extensive margin. All specifications are linear probability models, with Panel A being a level-level regression and Panel B level-log. Panel A has all its coefficients and standard errors multiplied by 10^5 for the sake of presentation. Interpretations is as follows: 100k more tweets in a PT municipality increases the probability of a protest to happen in 1.02%, while this effect if of 4.36% in municipalities without PT mayors.

All results keep stable, when comparing to table 5.4, except for the education subsets in Panel A, where low education has a higher effect of social media. However, this effect is only significant at the 10% level.

Table 5.5: Heterogeneous effects - extensive margin

	Protestors$_t$					
	High educ	Low educ	High income	Low income	PT	Not PT
			†Panel A: Level-Level Regressions			
Tweets$_t$	2.488**	4.706*	2.496**	0.375	1.022***	4.363*
	(1.027)	(2.728)	(1.035)	(0.489)	(0.092)	(2.570)
Tweets$_{t-1}$	−0.079	−0.537	−0.092	−0.154	1.420***	−1.240**
	(0.582)	(1.037)	(0.581)	(0.248)	(0.267)	(0.628)
Protestors$_{t-1}$	−0.230***	−13.376***	−0.230***	−11.840***	−0.788***	−0.147***
	(0.063)	(1.793)	(0.063)	(1.913)	(0.138)	(0.057)
F stat	3.69	1.11	3.7	0.88	2.28	3.85
p-value	(0.00)	(0.00)	(0.00)	(0.999)	(0.00)	(0.00)
R^2	0.179	0.060	0.180	0.049	0.193	0.168
Adjusted R^2	0.131	0.006	0.131	−0.007	0.108	0.124
			Panel B: Log-Log Regressions			
log(Tweets$_t$)	0.024***	0.002**	0.025***	0.0005	0.026***	0.016***
	(0.002)	(0.001)	(0.002)	(0.001)	(0.003)	(0.001)
log(Tweets$_{t-1}$)	0.007***	0.0002	0.006***	−0.0001	0.003	0.007***
	(0.002)	(0.0004)	(0.002)	(0.0003)	(0.003)	(0.001)
log(Protestors$_{t-1}$)	−0.009***	−0.009***	−0.009***	−0.008***	−0.017***	−0.005**
	(0.002)	(0.001)	(0.002)	(0.0005)	(0.003)	(0.003)
F stat	4.19	1.13	4.18	0.89	2.61	4.19
p-value	(0.00)	(0.00)	(0.00)	(0.999)	(0.00)	(0.00)
R^2	0.199	0.062	0.198	0.050	0.214	0.180
Adjusted R^2	0.151	0.007	0.151	−0.006	0.132	0.137
City FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
State*Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	43,824	43,872	43,848	43,848	10,896	76,800

Notes: Clustered standard errors, adjusted within micro-regions, in parenthesis. All logarithms were calculated adding 1 inside. *p<0.1; **p<0.05; ***p<0.01.
† All coefficients and standard errors in Panel A were multiplied by 10^5 for the sake of presentation.

What can also be seeing, again, is that results show contemporary activity still dominating past activity in terms of magnitude. This is a strong pattern that I will attempt to disentangle in the following chapter.

6
Analyzing dynamics and content of social media

Here I attempt to isolate the potential mechanisms driving the relationship between protests and social media. Having data on social media activity allows me to analyze dynamics and content shared through the network. The two mechanisms I investigate, behind the effect of social media on protests, are information diffusion and coordination. I disentangle the first into two types, information on the existence of a protest or information on the main target of the protests, *i.e.*, the government. The second I separate into three types, coordination between days, coordination hours before the protests or live coordination.

The higher, and systematic, importance of contemporary social media activity over past activity in the effects described in chapter 5 indicates two aspects of how social media is affecting the protests. If content shared through social media is, in fact, important only in the same day as the protest occurs, then (i) the coordination channel is active, and (ii) coordination is happening in a high frequency. People are deciding to attend a protest after an intra-day coordination, either inferring the number of potential protestors hours prior to the event or analyzing how violent the protest has become, for example. If this is true, results should hold if more lags are added to the regression.

To see if that is the case, I run the following regressions, adding up to five lags of protests and twitter activity and creating an accumulated variable for both:

$$Protestors_{ist} = \beta_1 Tweets_{ist} + \sum_{a=1}^{5} \gamma_a Tweets_{ist-a} + \sum_{a=1}^{5} \delta_a Protestors_{ist-a} + \\ + FE_i + FE_t + FE_s * FE_t + u_{ist} \tag{6-1}$$

$$Protestors_{ist} = \beta_1 Tweets_{ist} + \gamma \sum_{a=1}^{5} Tweets_{ist-a} + \delta \sum_{a=1}^{5} Protestors_{ist-a} + \\ + FE_i + FE_t + FE_s * FE_t + u_{ist} \tag{6-2}$$

Tables 6.1 and 6.2 shows the results in level-level and log-log specifications, respectively. Columns 1 through 4 show the results for regressions 6-1, while column 5 shows the result of regression 6-2.

Results are consistent across tables, with contemporary tweets having a significant and positive effect in the number of protestors. All specifications also have significant joint tests statistics.

The pattern of contemporary tweets dominating past activity still holds for the log-log specifications. However, it breaks for the level-level specifications in columns 2 and 3 of table 6.1, and reemerges for columns 4 and 5.

Given all the potential non-linearities not accounted for when adding all these lags, and the fact that the pattern is consistent across tables 5.1, 5.2, 5.3, 5.4, 5.5 and 6.2, indicate that what was really driving the protestors, through Twitter, was intra-day coordination, rather than inter-day coordination.

However, intra-day coordination could be happening because of two reasons. Information could be shared prior to the protests or during them. Both cases would relate to the idea of logistical coordination, the latter informing protestors already in the streets with "live coverage", giving updates on the level of violence and obstacles. The former would point to a pre-protest coordination, by informing on time and place and providing a more precise estimate on the expected number of protestors later on. To evaluate which was the case I analyze the content shared in each tweet, identifying shared information that related to live protest communication, *e.g.*, posts containing expressions such as "happening now" or "here on the streets". All expressions used to separate groups of tweets are presented in appendix section A.8.

Second, to inspect whether the informational channel was active I also analyze content shared between users. I look at tweets containing expressions related to information on scheduled protests and those with information on government decisions, institutions and politicians or government propaganda (good or bad). After selecting tweets containing such information, less than 1% of them contained instructions or announcements about scheduled protests. Therefore, I consider the informational channel through knowledge about the existence of protests to be irrelevant in this case. I continue by analyzing the informational channel only through relevant information on the main target of protests, the government.

Table 6.3 shows the results of these investigations, using the following panel regressions with the full set of city and time fixed effects, as well as state trends:

$$\text{Protestors}_{ist} = \beta_1 \text{Tweets with GovInfo}_{ist} + \beta_2 \text{Tweets without GovInfo}_{ist} + \\ + FE_i \ + \ FE_t + \ FE_s * FE_t \ + \ u_{ist} \tag{Twitter activity about the government}$$

$$\text{Protestors}_{ist} = \beta_1 \text{Tweets During Protests}_{ist} + \beta_2 \text{Tweets not During Protests}_{ist} + \\ + FE_i + FE_t + FE_s * FE_t + u_{ist}$$

(Twitter activity during a protest)

In the first regression I separate tweets into those that carry relevant information on the government and those which do not, and show the results in column 1. In the second regression I separate the tweets between those that were tweeted during the protests and those which were not, and show the results in column 2.

Tweets without government information have a significant and positive correlation with the number of protestors, while those containing information on the government have, statistically, no effect. This result is intuitive given the recurrent pattern of the above regressions, that contemporary social media activity dominates the relationship between social media and protests. For the informational channel to be active, through relevant information on the government, such information needs to be quite shocking to motivate people to protest on the same day that the information was released. Moreover, this information would need to be released in such relevance every period. If today's information on government's acts are marginal relative to yesterday's information, then past social media activity should not have such a smaller role than contemporary activity. Although there were plenty of criticisms during the period, there was no evidence of relevant information on the government being released every day during the protests.

Results for the timing of the tweets, column 2, show that only tweets that had not been tweeted during the protests have a significant, and positive, effect in the number of protestors. Aligned with the results that contemporary activity dominates past activity, described in tables 5.1, 5.2, 5.3, 5.4, 5.5 and 6.2, the results of table 6.3 lead to the conclusion that the main driver of protestors, through social media, was a pre-protest, same day coordination and that information diffusion did not play a role.

Table 6.1: Level-level regressions adding lags

	Protestors$_t$				
	(1)	(2)	(3)	(4)	(5)
Tweets$_t$	2.348***	1.889***	1.865***	18.301***	12.433***
	(0.854)	(0.519)	(0.500)	(3.699)	(2.882)
Tweets$_{t-1}$	−0.328	0.294	0.239	−2.669	
	(0.623)	(0.246)	(0.275)	(1.805)	
Tweets$_{t-2}$	0.985*	−0.765**	−0.847**	−3.044*	
	(0.540)	(0.376)	(0.421)	(1.728)	
Tweets$_{t-3}$		3.615**	4.054**	1.005	
		(1.607)	(1.784)	(1.100)	
Tweets$_{t-4}$			−0.558***	−1.467***	
			(0.191)	(0.177)	
Tweets$_{t-5}$				2.153***	
				(0.299)	
Protestors$_{t-1}$	−0.131***	−0.243***	−0.178***	−0.214***	
	(0.032)	(0.049)	(0.023)	(0.044)	
Protestors$_{t-2}$	−0.096	−0.106***	−0.096***	−0.112*	
	(0.059)	(0.023)	(0.014)	(0.061)	
Protestors$_{t-3}$		−0.120**	−0.142***	−0.051	
		(0.048)	(0.054)	(0.049)	
Protestors$_{t-4}$			−0.089**	−0.077***	
			(0.041)	(0.024)	
Protestors$_{t-5}$				−0.098***	
				(0.019)	
$\sum_{t=-1}^{-5} Tweets_t$					0.011
					(0.108)
$\sum_{t=-1}^{-5} Protestors_t$					−0.143***
					(0.015)
City FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
State*Time FE	Yes	Yes	Yes	Yes	Yes
F Stat	9.03	23.92	22.11	33.45	17.19
p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
R^2	0.325	0.592	0.623	0.784	0.650
Adjusted R^2	0.289	0.567	0.595	0.761	0.613
Observations	84,042	73,080	58,464	40,194	40,194

Notes: Clustered standard errors, adjusted within micro-regions, in parenthesis. *p<0.1; **p<0.05; ***p<0.01

Table 6.2: Log-Log regressions adding lags

	log(Protestors$_t$)				
	(1)	(2)	(3)	(4)	(5)
log(Tweets$_t$)	0.122***	0.113***	0.119***	0.233***	0.226***
	(0.010)	(0.009)	(0.011)	(0.020)	(0.019)
log(Tweets$_{t-1}$)	0.017***	0.005	0.002	0.031***	
	(0.007)	(0.006)	(0.007)	(0.012)	
log(Tweets$_{t-2}$)	0.080***	0.043***	0.026***	0.033***	
	(0.007)	(0.006)	(0.007)	(0.010)	
log(Tweets$_{t-3}$)		0.085***	0.049***	0.005	
		(0.008)	(0.008)	(0.008)	
log(Tweets$_{t-4}$)			0.102***	0.034***	
			(0.012)	(0.008)	
log(Tweets$_{t-5}$)				0.029***	
				(0.008)	
log(Protestors$_{t-1}$)	−0.064***	−0.084***	−0.121***	−0.221***	
	(0.018)	(0.018)	(0.016)	(0.020)	
log(Protestors$_{t-2}$)	−0.013	−0.031*	−0.074***	−0.162***	
	(0.017)	(0.017)	(0.015)	(0.018)	
log(Protestors$_{t-3}$)		−0.014	−0.056***	−0.151***	
		(0.017)	(0.015)	(0.017)	
log(Protestors$_{t-4}$)			−0.063***	−0.163***	
			(0.016)	(0.019)	
log(Protestors$_{t-5}$)				−0.159***	
				(0.016)	
$\sum_{t=-1}^{-5} log(Tweets_t)$					0.025***
					(0.004)
$\sum_{t=-1}^{-5} log(Protestors_t)$					−0.172***
					(0.012)
City FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
State*Time FE	Yes	Yes	Yes	Yes	Yes
F Stat	4.86	4.76	4.59	4.84	4.79
p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
R^2	0.206	0.224	0.256	0.345	0.342
Adjusted R^2	0.163	0.177	0.200	0.273	0.270
Observations	84,042	73,080	58,464	40,194	40,194

Notes: Clustered standard errors, adjusted within micro-regions, in parenthesis. All logarithms were calculated adding 1 inside. *p<0.1; **p<0.05; ***p<0.01

Table 6.3: Mechanisms - Government Information and Live Coverage

	$Protestors_t$				
	(1)	No. of tweets		(2)	No. of tweets
Without gov info	2.612***	718,802	Not during Protests	3.511**	755,766
	(0.887)			(1.681)	
With gov info	−5.542	61,274	During Protests	−32.450	24,310
	(4.822)			(41.206)	
City FE	Yes			Yes	
Time FE	Yes			Yes	
State*Time FE	Yes			Yes	
F Stat	5.27			5.08	
p-value	(0.00)	(0.00)			(0.00)
R^2	0.302			0.298	
Adjusted R^2	0.267			0.263	
Observations	91,350			91,350	

Notes: Clustered standard errors, adjusted within micro-regions, in parenthesis. Groups of tweets were selected through text mining (expressions used can be found in appendix subsection A.8). *p<0.1; **p<0.05; ***p<0.01

7
Conclusion

In this article I present evidence of a positive impact of social media in the Brazilian protests of June 2013. The innovations presented here are twofold: analyzing the dynamics of the social media-protests relationship through social media activity data and identifying an effect of social media on protests on an environment without systematic media censorship.

I conclude that the effect of social media in the Brazilian protests of June 2013 was carried out through intra-day pre-protest coordination, and that information diffusion had no role in increasing protest participation through social media.

One complementary explanation, that does not alter the above conclusions, could be that different medias might have had complementary roles. It might have been the case that other social medias were used for inter-day coordination, such as Facebook and Youtube, and information diffusion might had played a role in increasing participation through conventional media outlets, such as newspapers and televised news.

References

[1] SHIRKY, C.. **The net advantage**. Prospect Magazine, 165, 2009.

[2] DIAMOND, L.. **Liberation technology**. Journal of Democracy, 21(3):69–83, 2010.

[3] BRUNS, A.; HIGHFIELD, T. ; BURGESS, J.. **The arab spring and social media audiences english and arabic twitter users and their networks**. American Behavioral Scientist, 57(7):871–898, 2013.

[4] RASHED, F.. **Twitter post**. `https://twitter.com/FawazRashed/status/48882406010257408`, 2011. March 18, 2011, 8:04 PM.

[5] HOWARD, P. N.; AGARWAL, S. D. ; HUSSAIN, M. M.. **When do states disconnect their digital networks? regime responses to the political uses of social media**. The Communication Review, 14(3):216–232, 2011.

[6] BARBERÁ, P.; JOST, J. T.; NAGLER, J.; TUCKER, J. A. ; BONNEAU, R.. **Tweeting from left to right is online political communication more than an echo chamber?** Psychological science, p. 1531–42, 2015.

[7] ANTENUCCI, D.; CAFARELLA, M.; LEVENSTEIN, M.; RÉ, C. ; SHAPIRO, M. D.. **Using social media to measure labor market flows**. Working Paper 20010, National Bureau of Economic Research, March 2014.

[8] EDMOND, C.. **Information manipulation, coordination, and regime change**. The Review of Economic Studies, p. 1422–1458, 2013.

[9] LITTLE, A. T.. **Communication technology and protest**. The Journal of Politics, 78(1):152–166, 2016.

[10] SHAPIRO, J. N.; SIEGEL, D. A.. **Coordination and security how mobile communications affect insurgency**. Journal of Peace Research, p. 1–11, 2015.

[11] WILSON, C.; DUNN, A.. **The arab spring| digital media in the egyptian revolution: descriptive analysis from the tahrir data set**. International Journal of Communication, 5:1248–1272, 2011.

[12] STEINERT-THRELKELD, Z. C.; MOCANU, D.; VESPIGNANI, A. ; FOWLER, J.. **Online social networks and offline protest**. EPJ Data Science, 4(1):1–9, 2015.

[13] ENIKOLOPOV, R.; PETROVA, M. ; MAKARIN, A.. **Social media and protest participation: Evidence from russia**. Working paper, Available at SSRN: `https://ssrn.com/abstract=2696236`, 2016.

[14] TUFEKCI, Z.; WILSON, C.. **Social media and the decision to participate in political protest: Observations from tahrir square**. Journal of Communication, 62(2):363–379, 2012.

[15] ADAY, S.; FARRELL, H.; FREELON, D.; LYNCH, M.; SIDES, J. ; DEWAR, M.. **Watching from afar: Media consumption patterns around the arab spring**. American Behavioral Scientist, p. 899–919, 2013.

[16] EARL, J.; MCKEE HURWITZ, H.; MEJIA MESINAS, A.; TOLAN, M. ; ARLOTTI, A.. **This protest will be tweeted: Twitter and protest policing during the pittsburgh g20**. Information, Communication & Society, 16(4):459–478, 2013.

[17] BASTOS, M. T.; MERCEA, D. ; CHARPENTIER, A.. **Tents, tweets, and events: The interplay between ongoing protests and social media**. Journal of Communication, 65(2):320–350, 2015.

[18] ACEMOGLU, D.; HASSAN, T. A. ; TAHOUN, A.. **The power of the street: Evidence from egypt's arab spring**. Working Paper 20665, National Bureau of Economic Research, November 2014.

[19] KERN, H. L.; HAINMUELLER, J.. **Opium for the masses: How foreign media can stabilize authoritarian regimes**. Political Analysis, p. 377–399, 2009.

[20] SHAPIRO, J. N.; WEIDMANN, N. B.. **Is the phone mightier than the sword? cellphones and insurgent violence in iraq**. International Organization, 69(02):247–274, 2015.

[21] YANAGIZAWA-DROTT, D.. **Propaganda and conflict: Evidence from the rwandan genocide**. The Quarterly Journal of Economics, 129(4):1947–1994, 2014.

[22] CAMPANTE, F. R.; DURANTE, R. ; SOBBRIO, F.. **Politics 2.0: The multifaceted effect of broadband internet on political participation**. Working Paper 19029, National Bureau of Economic Research, May 2013.

[23] MINER, L.. **The unintended consequences of internet diffusion: Evidence from malaysia**. Journal of Public Economics, 132:66–78, 2015.

[24] MANACORDA, M.; TESEI, A.. **Liberation technology: mobile phones and political mobilization in africa**. Technical report, Centre for Economic Performance, London School of Economics and Political Science, 2016.

[25] ENIKOLOPOV, R.; PETROVA, M. ; ZHURAVSKAYA, E.. **Media and political persuasion: Evidence from russia**. The American Economic Review, 101(7):3253–3285, 2011.

[26] TATAGIBA, L.. **1984, 1992 e 2013. sobre ciclos de protestos e democracia no brasil/1984, 1992 and 2013. on protest cycles and democracy in brasil**. Política & Sociedade, 13(28):35, 2014.

[27] ALESSI, G.. **‘ o vandalismo afasta a população da causa’, dizem usuários do transporte público em são paulo.** https://noticias.uol.com.br/cotidiano/ultimas-noticias/2013/06/12/usuarios-de-transporte-publico-em-sp-torcem-pela-reducao-da-tarifa-mas-criticam-protestos-violentos.htm, June 2013. [Online; posted 12-June-2013]. Accessed 23-April-2017.

[28] PRESSE, F.. **Protestos contra aumento das passagens abalam são paulo e rio**. `http://g1.globo.com/mundo/noticia/2013/06/protestos-contra-aumento-das-passagens-abalam-sao-paulo-e-rio.html`, June 2013. [Online; posted 14-June-2013]. Accessed 23-April-2017.

[29] RIO DE PAZ. **Exigimos escolas, hospital e segurança no padrão Fifa**. `http://blogs.oglobo.globo.com/ancelmo/post/exigimos-escolas-hospital-seguranca-no-padrao-fifa-500401.html`, 2013.

[30] AGUILAR, A.; FERRAZ, C.. **What drives social unrest? evidence from brazil's protests**. Manuscript, 2014.

[31] ROCHA, A.. **Do Politics Respond to Political Protests? Effects of the 2013 Protests in Brazil**. Master's thesis, PUC-Rio, `http://www.econ.puc-rio.br/index.php/publicacao`, 2015.

[32] CARDOSO, G.; DI FATIMA, B.. **Movimento em rede e protestos no brasil qual gigante acordou?** Revista ECO-Pós, 16(2):143–176, 2013.

[33] PHELAN, O.; MCCARTHY, K. ; SMYTH, B.. **Using twitter to recommend real-time topical news**. In: PROCEEDINGS OF THE THIRD ACM CONFERENCE ON RECOMMENDER SYSTEMS, p. 385–388. ACM, 2009.

[34] HERMIDA, A.. **From tv to twitter: How ambient news became ambient journalism**. Media/Culture Journal, 13(2), 2010.

[35] IBOPE. **Avaliação das manifestações até o momento**. `http://g1.globo.com/brasil/noticia/2013/06/veja-integra-da-pesquisa-do-ibope-sobre-os-manifestantes.html`, June 2013.

[36] SANDERSON, E.; WINDMEIJER, F.. **A weak instrument -test in linear {IV} models with multiple endogenous variables**. Journal of Econometrics, 190(2):212 – 221, 2016. Endogeneity Problems in Econometrics.

A
Appendix

A.1
Twitter data

List of hashtags are shown in table A.1.

Table A.2 shows a sample of the information provided by GNIP data for a single tweet.

In table A.3 I show panel regressions of Twitter activity on main city characteristics to see what are the main determinants of twitter activity in a municipality. Since these are pure controlled correlations, a few of these variables may be capturing the effect of similar ones that appear as non significant.

A.2
Matching procedure

Brazilian cities are uniquely identified by name-state. There can be no two cities with the same name inside one state. Given that, I concatenate city name with state name, or state abbreviation, and run a partial matching algorithm to match these city-state strings, and any useful permutation of them, to IBGE data set on Brazilian cities in 2013.

For example, take the city of Campinas, located in the state of São Paulo. The partial matching algorithm will try to match the permutations of city name and state name or abbreviation - *CAMPINASSP, SPCAMPINAS, CAMPINASSAOPAULO, SAOPAULOCAMPINAS* - with the self-reported user location.

A.3
GPS only

Here I present the same tables as in the main body of the text, but using only tweets with GPS coordinates as my measure of social media activity. The goal of this exercise is to see whether GPS tweets are different than the tweets that can only be located through the user self-reported location.

Table A.4 shows the results for protest behavior. Table A.5 shows the results for the intensive and extensive margins. Tables A.6 and A.7 show the results for the addition of lags.

A.4 Instrumental variables

Even though I dispose of a high frequency panel, no set of fixed effects can deal with a reverse causality issue. Reverse causality, here, presents itself from the fact that, as the protests increase in size, they become more interesting to talk about in the social media. Simultaneously, higher social media activity might induce more people to streets.

Here I present an attempt to deal with reverse causality with an instrumental variables approach. For that, I will exploit differences in internet quality across states and internet access across municipalities as a potential source of exogenous variation in the number of tweets in a given municipality.

Under this approach I cannot employ the whole set of time and city fixed effects because internet access (from the 2010 IBGE Census) and quality (from May 2013 ANATEL report) data are time invariant. I interact both measures to construct an instrument that varies at the city level.

To take into account unobserved variables that might be correlated with Twitter activity, I control for a range of time invariant city characteristics. These controls cover demographics, infrastructure, political party preferences, education and income. Besides those, I also control for time and state fixed effects.

The identification assumption is that, conditional on city observable characteristics, state and day fixed effects, variations in internet access and quality across municipalities are orthogonal to unobserved determinants of protest participation.

The regressions to be estimated follow the standard IV setting:

$$Tweet_{ist} = \alpha Quality_s * access_i + \gamma X_i + FE_s + FE_t + FE_s * FE_t + e_{ist} \tag{1st stage}$$

$$Prot_{ist} = \beta \widehat{Tweet}_{ist} + \gamma X_i + FE_s + FE_t + \ FE_s * FE_t + u_{ist} \tag{2nd stage}$$

where i refers to the city, s to the state and t to time. X is the vector of time invariant control variables, $Tweet$ is the variable that accounts for the amount of tweets and $Prot$ accounts for the number of protestors. FE stands for fixed effects and e_{ist} and u_{ist} are the regression errors.

Table A.8 shows the results of the above regressions. Table A.9 adds the first lag of Twitter activity and number of protestors to the above equations. Tables A.10, A.11 and A.12 shows the effects for intensive margin, extensive margin and heterogeneous effects, respectively, as is shown in chapter 5. Panels A show level-level regressions and Panels B show log-log regressions. Interpretations follow the same pattern as the panel regressions presented in chapter 5. All weak instrument F statistics were calculated following (36).

One difference from the regressions in chapter 5 is that when clustering in micro-regions the instrument is no longer significant at the 5% threshold in the first stage for the level-level specifications (except for the extensive margin), however, they stay significant in the log-log specifications. Weak instrument tests are displayed, with its p-value below it, showing that the instrument is not weak the log-log specifications.

It can be seen from the results that treating reverse causality increases the coefficient of contemporary Twitter activity and barely changes the remaining coefficients. The contemporary effect increases in more than twice the value of the coefficient in panel regressions.

A.5
Count data approach

Here I present an estimation using a count environment. Dependent variable is, by nature, non-negative and discrete, *i.e.*, a count variable. The issue underlying this approach is that overdispersion is too big, rendering poisson, and even negative binomial estimates, unreliable (average number of protestors is 332 and standard deviation 4541).

Table A.13 shows the results for two specifications: model (1) fixed effects poisson and model (2) fixed effects negative binomial. The displayed coefficients show the incidental rate ratio. For the poisson model, 1000 more tweets generates 7% more people on the streets, on average.

A.6
Extensive margin

Table A.14 shows the extensive margin estimated through logit and fixed effects logit. Coefficients present the average marginal effect of an increase of $\Delta Tweet_{ij}$. For instance, a 5% increase in the number of tweets in a single day (36k) would increase the probability of a protest happening by 39% according to the fixed effects logit specification.

A.7
Hours disaggregation

Table A.15 shows panel regressions between protest attendance and Twitter activity disaggregated into moments of the day. Night tweets are classified as those which were posted between 6 pm and 11:59 pm, while non-night tweets are all the remaining tweets. It can be seen that both sets of tweets are significant, with non-night tweets having higher coefficients.

Drawing from the fact that contemporary tweets were the main force behind the effect of Twitter on the protests having both sets of tweets positive and significant suggests that there is no specific effect of time of the day on the protests. This comes from the fact that different cities had protests in different moments of the day.

A.8
Text mining

Expressions considered for tweets during protests are the following:

"ao vivo", "aovivo", "nesse instante", "neste instante", "na rua agora", "acontecendo agora", "nesse momento", "neste momento", "ha pouco", "pouco tempo atras", "violencia agora", "protestando agora", "estou na rua", "protesto acontecendo", "neste exato momento", "nesse exato momento", "nesse exato instante", "neste exato instante", "no momento", "no instante", "estamos aqui", "estou aqui", "por enquanto", "live", "protestando", "começando", "estamos na rua", "estamos no protesto", "estou no protesto", "estamos na manifestacao", "estou na manifestacao", "manifestacao acontecendo", "protesto agora", "manifestacao agora", "protesto ocorrendo", "manifestacao ocorrendo", "violencia policial acontencendo", "violencia policial agora", "quebra quebra acontencendo", "quebra quebra rolando", "quebra-quebra acontencendo", "quebra quebra", "agora pouco", "quebra-quebra rolando","violencia ocorrendo", "truculencia agora", "aqui na manifestacao", "aqui no protesto", "aqui na passeata"

Expressions considered for tweets containing information on the government are the following:

"corrupcao", "pronunciamentodilma", "foradilma", "fora dilma", "dilma", "gastospublicos", "gastos publicos", "contraacorrupcao", "naomerepresenta", "naonosrepresentam", "indignacao", "corrupto", "vergonha".

Table A.1: List of Hashtags

#pronunciamentodilma	#mudaBrasil	#protestoSE
#calabocadilma	#nãoéapenas20centavos	#protestoSP
#tamojuntodilma	#nãoésobre20centavos	#protestoSP
#dilmacadêobolsaf	#nãoésobrevintecentavos	#protestoTO
#20centavos	#nãosãoapenas20centavos	#protestoVIX
#vintecentavos	#naoesobre	#saimosdofacebook
#acordabrasil	#naoésobre	#semviolencia
#agoravai	#naomerepresenta	#SP17J
#amanhavaisermaior	#naonosrepresentam	#SP18J
#AnonymousBrasil	#nãonosrepresentam	#SP19J
#anonymousBR	#nãomerepresenta	#SP20J
#as5causas	#obrasilacordou	#SP21J
#causa	#occupySP	#SP22J
#causabrasil	#ogiganteacordo	#tarifazero
#contraoaumento	#ogiganteacordou	#vemprarua
#copapraquem	#opovoacordou	#vamosarua
#emprogresso	#passeata	#VerásQueUmFilhoTeuNãoFogeALuta
#essaeacausa	#passelivre	#mpl
#boicotecopa	#PasseLivre	#protestosbrasil
#foracopa	#protesto	#curagay
#indignacao	#protestoAC	#nãoacuragay
#manifestacao	#protestoAL	#urnaeletrônica
#manifestacaoAC	#protestoAM	#changebrasil
#manifestacaoAL	#protestoAP	#changebrazil
#manifestacaoAM	#protestoBA	#contraacorrupção
#manifestacaoAP	#protestoBH	#pelademocracia
#manifestacaoBA	#protestoBR	#ditadura
#manifestacaoBH	#protestobrasil	#democracia
#manifestacaoCE	#protestoCE	#códigoflorestal
#manifestacaoDF	#protestoDF	#reformapolítica
#manifestacaoES	#protestoES	#reformasjá
#manifestacaoGO	#protestoGO	#reformajá
#manifestacaoMA	#protestoMA	#acessibilidade
#manifestacaoMG	#protestoMG	#desburocratizacao
#manifestacaoMS	#protestoMS	#reformaagrária
#manifestacaoMT	#protestoMT	#atomédico
#manifestacaoPA	#protestoPA	#amanhãvaisermaior
#manifestacaoPB	#protestoPB	#revoltadovinagre
#manifestacaoPE	#protestoPE	#sempartido
#manifestacaoPI	#protestoPI	#reformatributária
#manifestacaoPOA	#protestoPOA	#monarquiajá
#manifestacaoPR	#protestoPR	#primaverabrasileira
#manifestacaoRIO	#protestoRIO	#aruaénossa
#manifestacaoRJ	#protestoRJ	#fimdovotoobrigatório
#manifestacaoRN	#protestoRN	#marcocivil
#manifestacaoRO	#protestoRO	#gastospúblicos
#manifestacaoRR	#protestoRR	#criaçãodenovospartidos
#manifestacaoRS	#protestoRS	#nãoesqueceremos
#manifestacaoSC	#protestos	#forarenan
#manifestacaoSE	#protestosbr	#grevedoscaminhoneiros
#manifestacaoSP	#protestosbr	#grevecaminhoneiros
#manifestacaoTO	#protestosBR	
#manifestacaoVIX	#protestoSC	

Table A.2: Sample of data generated by GNIP and Twitter

id	tag:search.twitter.com,2005:348492702582710275
verb	posted
posted time	22/06/2013
body	Indo pra Paulista, quem vai? #vemprarua
object	links and technical info
hashtags	vemprarua
retweetCount	0
language	pt
location	polygon of 4 coordinate points
geo	-23.66842281, -46.46119595
account ID	id:twitter.com:53913416
friends	146
followers	257
Given location	NA

Table A.3: Determinants of social media activity

	Tweets$_t$			
	(1)	(2)	(3)	(4)
Protests$_t$	0.101***	0.101***		
	(0.026)	(0.025)		
Urban	−0.390***	−0.390***	−0.649***	−0.649***
	(0.129)	(0.138)	(0.207)	(0.195)
Men	2.387	2.387**	3.627	3.627**
	(1.566)	(1.138)	(2.338)	(1.677)
High school	0.773*	0.773	1.601***	1.601**
	(0.426)	(0.619)	(0.401)	(0.659)
Electricity	0.040	0.040	0.115	0.115
	(0.093)	(0.120)	(0.144)	(0.191)
Sewer	−0.006	−0.006	−0.015	−0.015
	(0.019)	(0.031)	(0.026)	(0.047)
Water	0.255**	0.255	0.360**	0.360
	(0.116)	(0.168)	(0.176)	(0.261)
Cell Phones	0.034	0.034	0.045	0.045
	(0.038)	(0.090)	(0.055)	(0.133)
Log(GDP)	5.333	5.333**	7.027	7.027**
	(4.079)	(2.425)	(5.753)	(3.418)
Log(population)	16.272***	16.272***	25.134***	25.134***
	(5.119)	(4.717)	(8.406)	(7.229)
Age group 1	167.774	167.774**	190.889	190.889*
	(103.429)	(84.394)	(136.479)	(104.145)
Age group 2	167.280	167.280**	189.252	189.252*
	(103.826)	(85.031)	(136.707)	(104.729)
Age group 3	166.000	166.000**	187.538	187.538*
	(103.469)	(84.586)	(136.199)	(104.026)
Age group 4	169.067	169.067**	192.100	192.100*
	(104.108)	(84.896)	(137.269)	(104.669)
Observations	91,800	91,800	91,800	91,800
Clustering	City	State	City	State
Time FE	Yes	Yes	Yes	Yes
R^2	0.018	0.018	0.255	0.255
Adjusted R^2	0.018	0.018	0.254	0.254

Notes: All logarithms were calculated adding 1 inside. All variables from Urban until Cell phones represent the share of population or households that posses that specific characteristic. Age groups defined as the share of population that are in the cohorts 10-24, 25-39, 40-54, 60 and older. *p<0.1; **p<0.05; ***p<0.01.

Table A.4: GPS only - Protest Behavior

	Protestors_t			
	(1)	(2)	(3)	(4)
Panel A: Level-Level Regressions				
Tweets_t	40.560***	40.853***	48.305***	47.675***
	(12.112)	(14.004)	(16.586)	(17.949)
Tweets_{t-1}			−10.299	−10.175
			(7.431)	(8.162)
Protestors_{t-1}			−0.067**	−0.100***
			(0.027)	(0.030)
F stat	22102.35	10.36	7891.22	11.4
p-value	(0.00)	(0.00)	(0.00)	(0.00)
R^2	0.326	0.372	0.350	0.403
Adjusted R^2	0.326	0.336	0.350	0.367
Panel B: Log-Log Regressions				
$\log(\text{Tweets}_t)$	0.518***	0.502***	0.470***	0.486***
	(0.030)	(0.031)	(0.028)	(0.032)
$\log(\text{Tweets}_{t-1})$			0.064***	0.062***
			(0.016)	(0.019)
$\log(\text{Protestors}_{t-1})$			0.010	−0.080***
			(0.019)	(0.016)
F stat	7852.27	5.68	2527.88	5.66
p-value	(0.00)	(0.00)	(0.00)	(0.00)
R^2	0.146	0.245	0.147	0.251
Adjusted R^2	0.146	0.202	0.147	0.207
City FE	No	Yes	No	Yes
Time FE	No	Yes	No	Yes
State*Time FE	No	Yes	No	Yes
Observations	45,750	45,750	43,920	43,920

Notes: Clustered standard errors, adjusted within micro-regions, in parenthesis. All logarithms were calculated adding 1 inside. *p<0.1; **p<0.05; ***p<0.01

Table A.5: GPS only - Intensive and Extensive margins

	Protestors$_t$			
	Intensive margin		Extensive margin	
	Levels	Logs	Levels	Logs
Tweets$_t$	47.431**	1.026***	0.001**	0.057***
	(18.970)	(0.056)	(0.0002)	(0.004)
Tweets$_{t-1}$	−10.113	−0.193***	−0.0002*	0.008***
	(8.884)	(0.055)	(0.0001)	(0.003)
Protestors$_{t-1}$	−0.109***	−0.117***	−0.00001***	−0.010***
	(0.035)	(0.017)	(0.00001)	(0.002)
City FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
State*Time FE	Yes	Yes	Yes	Yes
Observations	8,184	8,184	43,920	43,920
F stat	6.47	4.84	3.74	4.6
p-value	(0.00)	(0.00)	(0.00)	(0.00)
R^2	0.464	0.392	0.181	0.214
Adjusted R^2	0.392	0.311	0.133	0.167

Notes: Clustered standard errors, adjusted within micro-regions, in parenthesis. All logarithms were calculated adding 1 inside. *p<0.1; **p<0.05; ***p<0.01

Table A.6: GPS only - Level-level regressions adding lags

	Protestors$_t$				
	(1)	(2)	(3)	(4)	(5)
Tweets$_t$	47.523***	40.320***	37.912***	181.810***	137.541***
	(16.388)	(11.718)	(10.721)	(48.033)	(36.692)
Tweets$_{t-1}$	−21.047**	−9.219	−8.407	−21.375	
	(10.430)	(6.043)	(6.591)	(15.302)	
Tweets$_{t-2}$	22.386***	−1.411	−4.314	−23.852	
	(8.243)	(2.535)	(3.892)	(17.648)	
Tweets$_{t-3}$		40.980**	51.790**	1.982	
		(19.647)	(23.820)	(11.739)	
Tweets$_{t-4}$			−12.751**	−11.003***	
			(6.436)	(2.558)	
Tweets$_{t-5}$				24.435***	
				(4.369)	
Protestors$_{t-1}$	−0.072*	−0.212***	−0.152***	−0.192***	
	(0.037)	(0.046)	(0.026)	(0.025)	
Protestors$_{t-2}$	−0.125**	−0.078***	−0.049***	−0.040	
	(0.049)	(0.016)	(0.018)	(0.083)	
Protestors$_{t-3}$		−0.126**	−0.177***	−0.038	
		(0.057)	(0.069)	(0.040)	
Protestors$_{t-4}$			−0.066***	−0.061***	
			(0.021)	(0.017)	
Protestors$_{t-5}$				−0.140***	
				(0.021)	
$\sum_{t=-1}^{-5} Tweets_t$					−0.884
					(2.094)
$\sum_{t=-1}^{-5} Protestors_t$					−0.108***
					(0.018)
City FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
State*Time FE	Yes	Yes	Yes	Yes	Yes
F Stat	12.99	19.43	18.71	31.17	19.31
p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
R^2	0.443	0.571	0.608	0.785	0.692
Adjusted R^2	0.409	0.542	0.576	0.760	0.657
Observations	42,090	36,600	29,280	20,130	20,130

Notes: Clustered standard errors, adjusted within micro-regions, in parenthesis. All logarithms were calculated adding 1 inside. *p<0.1; **p<0.05; ***p<0.01

Table A.7: GPS only - Log-Log regressions adding lags

	log(Protestors$_t$)				
	(1)	(2)	(3)	(4)	(5)
log(Tweets$_t$)	0.466***	0.444***	0.436***	0.736***	0.715***
	(0.032)	(0.031)	(0.033)	(0.059)	(0.057)
log(Tweets$_{t-1}$)	0.018	0.001	$-$0.005	0.033	
	(0.018)	(0.019)	(0.019)	(0.032)	
log(Tweets$_{t-2}$)	0.126***	0.059***	0.055**	0.031	
	(0.019)	(0.020)	(0.021)	(0.029)	
log(Tweets$_{t-3}$)		0.195***	0.172***	0.039	
		(0.025)	(0.025)	(0.025)	
log(Tweets$_{t-4}$)			0.138***	0.036	
			(0.026)	(0.027)	
log(Tweets$_{t-5}$)				0.038*	
				(0.021)	
log(Protestors$_{t-1}$)	$-$0.087***	$-$0.105***	$-$0.136***	$-$0.230***	
	(0.017)	(0.016)	(0.015)	(0.018)	
log(Protestors$_{t-2}$)	$-$0.029*	$-$0.045***	$-$0.081***	$-$0.157***	
	(0.016)	(0.016)	(0.015)	(0.017)	
log(Protestors$_{t-3}$)		$-$0.030*	$-$0.067***	$-$0.142***	
		(0.016)	(0.015)	(0.016)	
log(Protestors$_{t-4}$)			$-$0.068***	$-$0.154***	
			(0.016)	(0.017)	
log(Protestors$_{t-5}$)				$-$0.145***	
				(0.015)	
$\sum_{t=-1}^{-5} log(Tweets_t)$					0.034**
					(0.013)
$\sum_{t=-1}^{-5} log(Protestors_t)$					$-$0.167***
					(0.011)
City FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
State*Time FE	Yes	Yes	Yes	Yes	Yes
F Stat	5.6	5.45	5.03	5.42	5.34
p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
R^2	0.255	0.272	0.294	0.388	0.384
Adjusted R^2	0.210	0.222	0.236	0.316	0.312
Observations	42,090	36,600	29,280	20,130	20,130

Notes: Clustered standard errors, adjusted within micro-regions, in parenthesis. All logarithms were calculated adding 1 inside. *p<0.1; **p<0.05; ***p<0.01

Table A.8: IV regression - Protest Behavior

	Protestors$_t$				Tweets$_t$
	OLS	OLS	OLS	IV	1st stage
Panel A: Level-Level Regressions					
Tweets$_t$	2.404***	2.301***	2.380***	3.762***	
	(0.607)	(0.649)	(0.610)	(0.701)	
Quality*Internet					1.108
					(0.775)
F Statistic	29754.28	7.61	45.68	3.99	3.95
p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Weak Instrument F-stat					2.043
p-value					(0.153)
R^2	0.246	0.272	0.258		0.029
Adjusted R^2	0.246	0.236	0.252		0.022
Panel B: Log-Log Regressions					
log(Tweets$_t$)	0.179***	0.168***	0.200***	0.295***	
	(0.014)	(0.013)	(0.016)	(0.068)	
Quality*Internet					0.005***
					(0.001)
F Statistic	8756.32	4.57	18.49	10.64	145.05
p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Weak Instrument F-stat					43.085
p-value					(0.00)
R^2	0.087	0.183	0.123		0.524
Adjusted R^2	0.087	0.143	0.116		0.520
FE	No	City	State	State	State
Time FE	No	Yes	Yes	Yes	Yes
State*Time FE	No	Yes	Yes	Yes	Yes
Exogenous Controls	No	No	Yes	Yes	Yes
Observations	91,350	91,350	91,350	91,350	91,350

Notes: Clustered standard errors, adjusted within micro-regions, in parenthesis. All logarithms were calculated adding 1 inside. Exogenous controls consist of the following city variables (all gotten from the 2010 Census, unless specified otherwise): share of population that live in urban areas, that are men, that have a high school diploma or above; share of population that are in the cohorts 10-24, 25-39, 40-54, 60 and older; share of households that have electricity, that have sewer coverage, that have drinkable water access, that own at least one cellphone; log of population; log of the 2013 GDP (IBGE 2013); dummies for the party that won the executive seat at each city in 2012 for the 2013-2016 mandate (TSE 2012). $^*p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table A.9: IV - Protest Behavior with lags

	Protestors$_t$				Tweets$_t$
	OLS	OLS	OLS	IV	1st stage
	Panel A: Level-Level Regressions				
Tweets$_t$	2.473***	2.364***	2.452***	6.158***	
	(0.886)	(0.898)	(0.890)	(0.908)	
Quality*Internet					0.499
					(0.337)
Tweets$_{t-1}$	0.065	0.050	0.059	$-$2.289***	0.633***
	(0.390)	(0.416)	(0.393)	(0.680)	(0.009)
Protestors$_{t-1}$	$-$0.089***	$-$0.117***	$-$0.092***	$-$0.041	$-$0.014***
	(0.033)	(0.030)	(0.032)	(0.046)	(0.003)
F Statistic	9867.93	7.79	47.16	12.83	80.7
p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Weak Instrument F-stat					2.186
p-value					(0.139)
R^2	0.252	0.284	0.265		0.381
Adjusted R^2	0.252	0.248	0.259		0.377
	Panel B: Log-Log Regressions				
log(Tweets$_t$)	0.130***	0.139***	0.156***	0.383**	
	(0.010)	(0.010)	(0.012)	(0.184)	
Quality*Internet					0.002***
					(0.0002)
log(Tweets$_{t-1}$)	0.055***	0.056***	0.058***	$-$0.097	0.680***
	(0.006)	(0.008)	(0.008)	(0.124)	(0.011)
log(Protestors$_{t-1}$)	0.045**	$-$0.050***	0.022	0.019	0.010
	(0.023)	(0.018)	(0.022)	(0.021)	(0.006)
F Statistic	2994.81	4.56	18.91	15.79	388.25
p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Weak Instrument F-stat					49.318
p-value					(0.00)
R^2	0.093	0.189	0.126		0.748
Adjusted R^2	0.093	0.147	0.120		0.746
FE	No	City	State	State	State
Time FE	No	Yes	Yes	Yes	Yes
State*Time FE	No	Yes	Yes	Yes	Yes
Exogenous controls	No	No	Yes	Yes	Yes
Observations	87,696	87,696	87,696	87,696	87,696

Notes: Clustered standard errors, adjusted within micro-regions, in parenthesis. All logarithms were calculated adding 1 inside. Exogenous controls are the same as in table A.8. *p<0.1; **p<0.05; ***p<0.01

Table A.10: IV - Intensive margin

	Protestors$_t$				Tweets$_t$
	OLS	OLS	OLS	IV	1st stage
	Panel A: Level-Level Regressions				
Tweets$_t$	2.467***	2.343**	2.406***	5.937***	
	(0.874)	(0.914)	(0.908)	(1.109)	
Quality*Internet					1.681
					(1.064)
Tweets$_{t-1}$	0.062	0.064	0.062	−2.120***	0.616***
	(0.393)	(0.441)	(0.420)	(0.797)	(0.008)
Protestants$_{t-1}$	−0.091***	−0.128***	−0.108***	−0.052	−0.016***
	(0.033)	(0.031)	(0.032)	(0.048)	(0.003)
F Statistic	910.52	4.33	6.04	2.39	7.94
p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Weak Instrument F-stat					2.499
p-value					(0.114)
R^2	0.250	0.360	0.346		0.410
Adjusted R^2	0.250	0.277	0.288		0.358
	Panel B: Log-Log Regressions				
log(Tweets$_t$)	0.332***	0.593***	0.609***	1.042*	
	(0.025)	(0.039)	(0.041)	(0.603)	
Quality*Internet					0.001***
					(0.0004)
log(Tweets$_{t-1}$)	0.033	−0.106***	−0.112***	−0.418	0.701***
	(0.023)	(0.036)	(0.036)	(0.425)	(0.015)
log(Protestants$_{t-1}$)	−0.033	−0.110***	−0.077***	−0.062**	−0.033***
	(0.021)	(0.018)	(0.021)	(0.032)	(0.006)
F Statistic	400.01	3.98	5.26	4.52	105.92
p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Weak Instrument F-stat					10.701
p-value					(0.001)
R^2	0.128	0.341	0.315		0.903
Adjusted R^2	0.128	0.255	0.255		0.894
FE	No	City	State	State	State
Time FE	No	Yes	Yes	Yes	Yes
State*Time FE	No	Yes	Yes	Yes	Yes
Exogenous Controls	No	No	Yes	Yes	Yes
Observations	8,184	8,184	8,184	8,184	8,184

Notes: Clustered standard errors, adjusted within micro-regions, in parenthesis. All logarithms were calculated adding 1 inside. Exogenous controls are the same as in table A.8. *p<0.1; **p<0.05; ***p<0.01

Table A.11: IV - Extensive margin

	Protestors_t				Tweets_t
	OLS	OLS	OLS	IV	1st stage
Panel A: Level-Level Regressions					
Tweets_t	0.000041***	0.000026**	0.000036***	0.000280**	
	(0.000013)	(0.000011)	(0.000012)	(0.000115)	
Quality*Internet					0.499
					(0.337)
Tweets_{t-1}	0.000009***	0.000001	0.000007*	−0.000148**	0.633***
	(0.000003)	(0.000005)	(0.000004)	(0.000068)	(0.009)
Protest_{t-1}	−0.000000	−0.000002***	−0.000001	0.000002	−0.014***
	(0.000001)	(0.000001)	(0.000001)	(0.000002)	(0.003)
F Statistic	863.15	3.61	12.75	7.43	80.7
p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Weak Instrument F-stat					2.186
p-value					(0.139)
R^2	0.028	0.155	0.088		0.381
Adjusted R^2	0.028	0.112	0.081		0.376
Panel B: Level-Log Regressions					
$\log(\text{Tweets}_t)$	0.017***	0.018***	0.02***	0.037	
	(0.001)	(0.001)	(0.001)	(0.023)	
Quality*Internet					0.002***
					(0.0002)
$\log(\text{Tweets}_{t-1})$	0.008***	0.007***	0.007***	−0.004	0.68***
	(0.001)	(0.001)	(0.001)	(0.015)	(0.010)
$\log(\text{Protestors}_{t-1})$	0.006**	−0.006***	0.003	0.002	0.009
	(0.003)	(0.002)	(0.002)	(0.002)	(0.006)
F Statistic	2503.36	4.05	16.51	14.43	388.25
p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Weak Instrument F-stat					49.318
p-value					(0.00)
R^2	0.078	0.171	0.112		0.747
Adjusted R^2	0.078	0.128	0.105		0.745
FE	No	City	State	State	State
Time FE	No	Yes	Yes	Yes	Yes
State*Time FE	No	Yes	Yes	Yes	Yes
Exogenous Controls	No	No	Yes	Yes	Yes
Observations	87,696	87,696	87,696	87,696	87,696

Clustered standard errors, adjusted within micro-regions, in parenthesis. All logarithms were calculated adding 1 inside. Exogenous controls are the same as in table A.8. $\log(\text{Protestors}_{t-1})$ is the log of the amount of protestors in last period. Coefficients interpretations in the first 4 columns of Panel A are, *e.g.*, 100,000 more tweets increase the probability of an event happening, on average, by 12.87%, in the IV specification. *p<0.1; **p<0.05; ***p<0.01.

Table A.12: IV - Heterogeneous effects

	Protestors$_t$					
	High educ	Low educ	High income	Low income	PT	Not PT
	Panel A: Level-Level Regressions					
Tweets$_t$	5.843***	−3.604	6.127***	−2.470	2.120	6.191***
	(0.635)	(2.258)	(0.835)	(1.768)	(2.902)	(0.772)
Tweets$_{t-1}$	−2.089***	1.506	−2.263***	1.288	0.305	−2.103***
	(0.532)	(0.971)	(0.651)	(0.935)	(1.903)	(0.551)
F Statistic	7.35	0.38	6.89	0.04	10.63	12.74
p-value	(0.00)	(0.998)	(0.00)	(0.997)	(0.00)	(0.00)
	Panel B: Log-Log Regressions					
log(Tweets$_t$)	0.427***	−0.380**	0.467***	−0.258*	−0.064	0.465***
	(0.161)	(0.175)	(0.131)	(0.142)	(0.318)	(0.136)
log(Tweets$_{t-1}$)	−0.099	0.165**	−0.116	0.110*	0.233	−0.154
	(0.118)	(0.077)	(0.089)	(0.062)	(0.238)	(0.094)
F Statistic	9.83	0.66	9.77	0.29	2.69	13.29
p-value	(0.00)	(0.998)	(0.00)	(0.997)	(0.00)	(0.00)
FE	State	State	State	State	State	State
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
State*Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Exogenous Cont.	Yes	Yes	Yes	Yes	Yes	Yes
Observations	43,824	43,872	43,848	43,848	10,896	76,800

Notes: Clustered standard errors, adjusted within micro-regions, in parenthesis. All logarithms were calculated adding 1 inside. Exogenous controls are the same as in table A.8. Protestors in *t-1* omitted *p<0.1; **p<0.05; ***p<0.01

Table A.13: Count Data Specifications

	Incidental rate ratio: *Protest*$_t$	
	Poisson (1)	Negative Binomial (2)
Tweets$_t$	1.000069 ***	1.000119 ***
	(0.00001)	(0.00002)
City FE	Yes	Yes
Observations	8,525	8,525

Notes: Coefficients display the incidental rate ratio. Standard errors for the poisson model are clustered within cities. Standard error for the negative binomial are the standard MLE errors. *p<0.1; **p<0.05; ***p<0.01

Table A.14: Etensive margin - Average Marginal Effects of Logit Estimation

	$Protest_t$	
	Logit	FE Logit
	(1)	(2)
$Tweets_t$	$5*10^{-5}$ ***	$1*10^{-5}$ ***
	(0.000005)	(0.00002)
City FE	No	Yes
Observations	8,525	8,525

Notes: Coefficients display the average marginal/partial effects. City fixed effects included. Standard errors calculated through Delta method. *p<0.1; **p<0.05; ***p<0.01

Table A.15: Day vs Night Tweets

	$Protests_t$	
	Level-Level	Log-Log
$non\_night\_tweets_t$	4.281***	0.162***
	(0.988)	(0.019)
$night\_tweets_t$	1.774***	0.109***
	(0.639)	(0.010)
FE	City	City
Time FE	Yes	Yes
State*Time FE	Yes	Yes
Observations	91,800	91,800
R^2	0.276	0.197
Adjusted R^2	0.241	0.157

Notes: Clustered standard errors, adjusted within states, in parenthesis. All logarithms were calculated adding 1 inside. 5.1. *p<0.1; **p<0.05; ***p<0.01