Can chants in the street change politics' tune? Evidence from the 15M movement in Spain

Annalí Casanueva Artís*

November 13, 2022

Click here for the latest version

Abstract

What are the long term effects of protests? This paper studies how the level of attendance at simultaneous marches organized by the *15M* (the Spanish Occupy movement) impacted electoral behavior and political attitudes in the following decade. Using regional variation in weather shocks as an instrumental variable for the level of attendance at simultaneous marches, I find that cities with higher attendance are more concerned about corruption and vote more for left wing and anti-corruption parties and less for far-right parties. Using novel data from Twitter, I document, for the first time, a higher uptake of social media platforms after an offline protest and a persistent difference in online activity in cities with higher attendance. Using survey data, I also show a higher and longer-lasting electoral effect for people that have a social media account. Overall, this paper shows that street protests can have long-lasting effects on political concerns and electoral choices, explained, in part, by the creation of a persistent online social network.

Keywords: Protest, elections, networks, social media, corruption, Occupy movements, 15M
JEL code: D72, L82, L86

---

*Paris School of Economics. This work was partly funded by the Mexican Council of Science and Technology, CONACYT, to which I am grateful. I deeply thank Ekaterina Zhuravskaya for her supervision and support. I am in debt to Vladimir Avetian, François Bourguignon, Denis Cogneau, Matthew Gentzkow, Guillaume Geoffroy, Rubén Durante, Ruben Enikolopov, Sylvie Lambert, Jonathan Lehne, Ro'ee Levy, François Libois, Kenji Maillard, David Margolis, Hillel Rapoport, Cecile Rotner, Raúl Sánchez de la Sierra, Sulin Sardoschau, Kritika Saxena, Carlo Schwarz, Marco Tabellini, Oliver Vanden Eynde, Ambre Williams, Carlos Winograd, Liam Wren Lewis, Noam Yuchtman, Théo Zimmerman and Maiting Zhuang for providing valuable feedback. All errors remain my own.

1 Introduction

Protests against the *status quo* political systems have increased in recent years, reaching unprecedented frequency, size and scope (Brannen et al., 2020). From the Arab Spring and Occupy Wall Street to the Umbrella Revolution in Hong Kong and the feminist movement in Iran, protests have been rising in all regions of the world, at a world average of more than 11% per year in the last decade (Della Porta and Mattoni, 2015; Brannen et al., 2020). What are the effects of protests? Do they lead to a real political change?

Modern social movements differ from traditional protests in various ways which may impact their ability to convert mobilization into long-lasting change. While they have taken advantage of modern technologies to overcome the collective action problem and achieve the mobilization of millions of people (Enikolopov et al., 2020), they are often organized by people with little or no political experience who may lack the knowledge to turn successful mobilization into long-term political effects (Della Porta and Mattoni, 2015). Yet the empirical literature studying the causal effects of modern protests has so far only considered their short term effects.

This paper fills this gap by showing *long-term* effects of protests, and additionally brings forward, tests and documents a new possible mechanism that may explain part of these lasting results: the creation of an online social network as a result of street protests.

I use the 2011 *Indignados* or 15M[1] movement in Spain to study the short, medium and long term effects of street marches on electoral choices and political attitudes. The 15M movement is a modern social movement that is part of the wider family of "Occupy movements", characterized by the engagement of people without previous political experience, a heavy use of social media, and occupation of the public space (Della Porta and Mattoni, 2015). In the context of a difficult economic situation and a high salience of political corruption scandals, participants protested against corruption among politicians, criticized the political and economic system, and asked for increased citizen participation in politics (Fominaya and Cox, 2013).[2]

The characteristics of the 15M and the simultaneous marches they organized in a single day all over Spain make them well suited to studying the effects of modern protests. In contrast with other Occupy movements such as Occupy Wall Street, the protests occurred all over Spain, providing variation in the level of attendance that can be exploited to estimate causal effects. Additionally, the simultaneity of the protests reduces the risk that the success of a demonstration in one city could affect the success of demonstrations elsewhere.

Beyond its advantages for causal identification, the 15M took place in a period where the social media market was not yet saturated, i.e. there still were many potential social media users that hadn't yet signed up. This makes it possible to analyse the effects of

[1] The movement was named after the date of their first demonstration on the 15th May 2011.

[2] Unemployment rates were high (around 21% of the general population according to INE) and corruption scandals affecting high-positioned politicians were regularly reported in the media.

protests on both the intensive and extensive margin of social media use. Through the particular context of Spain, I can examine the effects of protests on attitudes toward corruption. Corruption has been shown to have long-term impacts on institutions (Solé-Ollé and Sorribas-Navarro, 2018; Dimant and Tosato, 2018).

To study the causal effect of different levels of participation in demonstrations, I exploit cross-city variation in weather shocks on the day of the march. If a city experienced either unusually rainy or unusually hot weather, fewer people will have attended the demonstration. Indeed, unpleasant weather has been linked to a decrease in outdoor activities (Bélanger et al., 2009; Tucker and Gilliland, 2007) including those related to political participation (Artés, 2014; Zhang, 2016; Madestam et al., 2013). In the context of 15M, I find that negative weather shocks significantly reduced attendance at protests by around 50%, consistent with media reports about the need for medical attention of some participants due to high temperatures.

Weather conditions have been linked to other variables that can also affect electoral outcomes, such as economic activity, health or migration (Fontenla et al., 2019; Acevedo et al., 2020; Mellon, 2021). Two elements of the identification strategy greatly help reduce concerns about the exclusion restriction. First, I only use weather shocks during the precise day of the march, which took place five months before the earliest elections considered. This already tackles a wide range of concerns, such as the direct effect of the weather on election day on electoral results (Arnold and Freier, 2016). Second, I control for the probability of the city having unpleasant weather on a random day in June. This control allows me to remove the potential correlation between mean city weather and characteristics impacting elections. I can thus interpret the instrument as measuring *weather shocks*. The weather in the days and weeks around the day of the demonstration is likely to be correlated with weather on the day of the demonstration and could affect other variables linked to electoral outcomes. Therefore, I include weather conditions during the other days of June 2011 (excluding 3 days before and 3 days after the the day of the demonstrations) as a control to show that the results are not driven by them.

In addition to these two elements, I conduct several supplementary checks that further mitigate possible remaining concerns about the exclusion restriction. First, I include weather conditions in the months surrounding the demonstrations in addition to the conditions in June 2011 and show that the results, again, are not driven by longer-term weather conditions. Second, sensitivity analysis (Cinelli and Hazlett, 2020b) shows that even if there still existed leftover correlation between my instrument and variables linked to electoral results, it is very unlikely that those correlations could overturn the results I found. Moreover, I consider concerns about spurious correlation between weather shocks and political preferences. I show that, on average, weather shocks on random days cannot predict protest attendance. I also show that weather shocks are conditionally uncorrelated with previous political outcomes and with contemporaneous unemployment rates, which themselves are correlated with protests and electoral behavior (Rosenstone, 1982; Stein, 1990; Nam, 2007; Burden and Wichowsky, 2014).

With this identification strategy, I estimate the effect of the differential level of attendance at protests (explained by weather conditions) on the electoral choices in all subsequent congressional and European elections up to 2019.[3][4] Using the full available period, I can capture the effects of protests in three different contexts. First, the analysis of the full sample gives us the big picture, with the advantage of having more precise estimates due to a higher number of observations. Second, by focusing on the elections after 2014, I capture the effects once the political offer includes new political parties that are connected with the *Indignados* movement and were not present in 2011. Finally, by restricting the sample to the last available election year (2019), I can analyze the effects of protests after the rise of far-right parties in the media and the public debate and test the long-term persistence of the impact of protests.

Results show that ten additional demonstrators lead to around 8 additional votes for left-wing parties, 6 additional votes for anti-corruption parties, and 2 votes less for far-right parties in each of the following elections. The effect is very similar for the different sub-samples that restrict the time frame to elections with new political options. Remarkably, the effects persist over time: the estimates restricted to the 2019 elections are not statistically different from those using the whole sample, showing long-term effects. Additionally, using answers on previous votes from post-electoral surveys, I show that the effect of demonstrations is not limited to the left: I observe that part of the reduction of the far-right vote in cities with larger demonstrations can be explained by a reduced transfer of voters from the right to the far-right. This suggests that left-leaning street protests can have effects on voters over the whole political spectrum.

I complement the city-level electoral results with individual-level data. Results coming from survey data show that higher attendance at the protest in a city causes a persistent increase in corruption concerns and in the probability that people declare being on the left half of the political spectrum. Ten additional demonstrators increase the number of people highly concerned about corruption by 3 and the number of people self-identifying as left-leaning by 11. Similarly to results coming from elections, higher protest attendance decreases the probability that people consider themselves at the far-right of the political spectrum. Those effects are still present even when considering only surveys conducted in 2019, confirming the presence of long-term effects.

Heterogeneity analyses on both the city and individual data sets show that even if people from different cities and with different characteristics react differently to protests, they all shifted leftwards as a consequence of higher protest attendance in their cities. In particular, right-leaning cities decreased their vote for far-right parties. People who have

[3] I focus on congressional and European elections because they are the two electoral events that are the most homogeneous at the national level, with campaigns focusing on nation-wide issues and similar parties running in all electoral districts. In addition, the 2014 European elections were the first elections for which the two parties closest to the *15M* movement –Podemos and Partido X– ran.

[4] 2019 is the latest election year at the time of writing.

post-secondary diplomas or were young when the marches happened are more likely to self-identify as being left-leaning and less likely to identify as right-leaning as a result of higher protest attendance in their cities.

The long-lasting effects of higher levels of attendance at a single street protest may seem puzzling. It is difficult to argue that three hours of marching could cause electoral changes almost a decade later out of thin air. Rather, I argue that protests have the capacity to create a lasting network that exposes its members to content related to the movement's views over time.

Modern social movements share an up-to-date knowledge about socialization dynamics on the Internet, through which they can self-mediatize (Della Porta and Mattoni, 2015; Castells, 2015). The 15M movement was no exception (Martínez Roldán, 2011). It successfully drew attention to the protests on social media: 15M became the most popular trending topic globally on Twitter for several days during the heyday of the simultaneous marches.[5] Additionally, survey data on protest participants documents that around a third of respondents started to use Twitter due to the 15M movement (Redes, Movimientos y Tecnopolítica, 2014).

I use geo-located Twitter data to argue that 15M street protests were successful in creating an online network that exposes people to a content close to the opinions of 15M in a persistent way. I proceed in several steps. First I create an index of social media activity using the first principal component (PCA) of three measures of Twitter usage: i) number of tweets, ii) number of distinct active users, and iii) new accounts created. I show that in cities with higher protest attendance, more Twitter accounts were created, and the general Twitter activity was higher in the month following the demonstration, while there was no difference in activity before the 15M. Then I show that users in those cities tweeted more about 15M related topics. Using state of the art sentiment analysis algorithms (Liu et al., 2019), I also show that cities with higher protests attendance produce a higher number of positive tweets about 15M. Third, I show that this regional difference in activity persists over time. In particular, people from cities with higher protest attendance still tweet more about 15M and tweet more about the main left-leaning party linked to the 15M movement (Podemos) before each congressional election.[6] Analysis of both Twitter and survey data shows that, after each major corruption scandal, people were more concerned about corruption and tweeted more about it in regions with higher protest attendance. Finally, I turn to survey data and show that the effects of protest attendance on electoral choices are stronger for people with social media accounts.

Other mechanisms may be at play. For example, voters can "vote with their feet" and decide to move to cities with political preferences closer to theirs (Tiebout, 1956; Hirschman, 1970). In that case, part of the long-term effects could come from an increase in left-wing voters in cities with higher number of protesters. Using data on migration

[5]This has been reported for example in Diario Sur or El Pais.

[6]See section 2.3 for a brief description of the party and its creation.

flows and a two way fixed effects strategy, I show that different weather conditions during the day of the marches between different cities cannot predict internal migration between theses cities.

Electoral inertia could be another channel explaining part of the long-term persistence of the results. Indeed, it is well known that parties obtaining good results in one election tend to obtain better results in subsequent elections as a consequence (Mullainathan and Washington, 2009; Erikson, 1971; Curto-Grau et al., 2012). I use voting history data from post-electoral surveys to show that, due to pleasant weather conditions during the demonstration, voters continue switching to parties defending positions close to 15M even years after the demonstrations.

In addition, protests could have effects through offline social networks or could have caused left-leaning parties to focus their campaign efforts in cities where protests have been more successful. In the last part of the mechanisms section, I discuss these alternative channels further and argue that it is unlikely that theses channels could fully explain the long-term persistence of the results.

This paper contributes to different strands of the literature. First, it contributes to the recent literature on the effects of protests (Madestam et al., 2013; Wasow, 2020; González, 2017; El-Mallakh et al., 2018; Mazumder, 2018, 2019; Brox and Krieger, 2021; Campbell, 2021; Larreboure and González, 2021; Levy and Mattsson, 2021) by analyzing not just the short-term but also the medium and long-term effects of protests.

The literature studying the political effects of protests in developed countries finds that protests tend to succeed in creating sympathy among people outside the movement, increase concern about the problems that the movement raises, persuade individuals towards the protesters' views and change their electoral behavior and attitudes, helping empower historically underrepresented groups, at least in the short run (Madestam et al., 2013; González, 2017; Wasow, 2020; Mazumder, 2018, 2019; Klein Teeselink and Melios, 2021; Sisco et al., 2021; Campbell, 2021; Skoy, 2021; Larreboure and González, 2021; Levy and Mattsson, 2021). My results confirm these findings and show, additionally, that those effects can be long-lasting. In particular, I show that a march in 2011 can change electoral results in 2019, change self-placement on the left-right political axis, and change the level of concern about corruption (particularly after corruption scandals) at least up to almost a decade after the march took place. I also provide new evidence on the demographic characteristics of individuals that experience the strongest effects of protest attendance. Building on Mazumder (2019), I find that not only younger age, but also a higher educational level results in stronger effects in the change in political attitudes as a result of protests.

The analysis I conduct on the role of social media in explaining the long term effects of protest relates to the literature analyzing both the determinants and effects of social media uptake and activity.

First, to the best of my knowledge, this paper is the first to provide evidence of the creation of an online network following offline protest. The causes of social media uptake have not been been studied by themselves in the empirical economic literature. The uptake of social media has mainly been considered as a treatment whose effects on social outcomes is the matter of interest. As such, several papers exploit different instruments that exogenously explain part of the uptake (Müller and Schwarz, 2020; Enikolopov et al., 2020). In this context, the analysis of the causes of the uptake is partial as it is not central to the study. In contrast, I directly show, causally, that higher attendance at street protests can induce the creation of social media accounts. In particular, I show: i) that higher attendance at protests increases social media uptake; ii) that the social media activity is different in cities with higher attendance: they tweet more about 15M (both in terms of number of tweets and in the number of distinct users), about Podemos (a new political party born in the aftermath of 15M that addresses most of their demands) and about corruption; iii) that this difference in online activity across cities still existed at least up to 2019; and iv) that the effect of protest size is stronger for people with a social media account. Overall this provides evidence suggesting that protests can have an effect not just by sending an informational signal (Lohmann, 1993), but also through a particular use of social media networks.

Second, this paper also contributes to the literature on the effects of social media (Enikolopov et al., 2020; Rotesi, 2019) and other networking technologies on political participation (Amorim, 2016; Manacorda and Tesei, 2016; Campante et al., 2017).[7] In general, the literature suggests that online social networks have an impact on electoral behavior. Accordingly, I show that the social media network spawned around the 15M is linked to a change in electoral results. Similarly to what I argue in this paper, Campante et al. (2017) show that new emerging political actors are able to take advantage of new technologies to attract disenchanted or demobilized voters, and that these new forms of mobilization eventually feed back into the mainstream electoral process, converting "exit" back into "voice".

Empirical social media scholars have also provided evidence supporting (Allcott et al., 2020; Levy, 2021) or contradicting (Barberá, 2014; Boxell et al., 2017) an increase in political polarization as an effect of social media.[8] The heterogeneity analysis on the effects of protest attendance conducted in this paper suggest a leftward shift for people from the whole political spectrum and does not find any evidence on a possible increase in polarization. Overall, I find that not all uses of social media lead to polarization. Social media activity resulting from grass-root street marches can shift political preferences for the whole political spectrum toward mirroring the movement's opinions.

The rest of the paper proceeds as follows: section 2 presents the historical and national context, and describes the movement; section 3 describes the data; section 4 describes and discusses the empirical strategy; section 5 discusses the results; section 6 studies the role

[7]See Zhuravskaya et al. (2019) for a review.
[8]See Zhuravskaya et al. (2019) for a review.

of social media in explaining the results; section 7 discusses alternative mechanisms; and section 8 concludes.

2 Background and description of the movement

2.1 Historical Context

The democratic transition after the end of Franco's regime in the eighties led to a political system (and generally public institutions) that benefited the existing elites (Monterde, 2015) with an electoral law that disadvantages left-wing parties and a lack of left-wing media. In the late 2000s, the social democrats (PSOE, Partido Socialista Obrero Español) adopted public austerity measures following European recommendations even though per capita social expenditure was already lower than the average in the 15-member European Union (Navarro, 2011). In addition, several corruption scandals affecting political parties and the Crown came in the spotlight starting in 2007, with new evidence appearing in the news regularly. In 2011, Spain was in a deep economic crisis due in part to the crash of a housing bubble which led to numerous families losing their homes. The unemployment rate was around 25% and youth unemployment greater than 40%. At the same time, banks were being bailed out with the State's money.

In response to this situation, several social movements appeared, highlighting specific problems (e.g. "*V de Vivienda*", V for Housing in English) or simply representing a concrete sector of the population (Youth Without Future or Movement of Mortgage Victims). The 15M movement embraced these previous social movements and created a more holistic movement both in terms of demands and in terms of sympathizers. The movement had very high support from the general population. Three out of four people agreed with their demands and one out of two with its strategy (Sampedro and Lobera, 2014).

2.2 The 15M movement

2.2.1 Description and categorization

The 15M movement (named after the date of their first demonstration, on May 15, 2011), also called "Los Indignados" (The Outraged) or "Spanish Revolution", was a social movement that protested against the political and economic institutions. Participants asked for change in the political system and more participation from citizens. The movement was assembly-based and leader-less. It adopted very effective strategies to become highly represented in online social networks that allowed it to be "self-mediatized", skipping the filters of traditional media (Castells, 2015).

The Indignados movement is part of a wave of movements of protest and dissent against politicians that started in the late 2000s. These movements have been labelled in various forms such as “Occupy movements” or “Movements of the Crises” (Della Porta and Mattoni, 2015); and also fit in the broader description of “New Social Movements” (Buechler, 1995), “Intelligent Masses” (Martínez Roldán, 2011) or “Smart Mobs” (Rheingold, 2002). Even if the particular context of each individual movement is unique they share some common features:

1. They all involved large numbers of protesters who appeared to be independent of the usual political actors (unions or traditional political parties) and included some grassroots groups (Della Porta and Mattoni, 2015). In the case of the Indignados movement, it included several grassroots organizations devoted to more specific problems such as *V de Vivienda* (V of Housing) or *Youth without future*. Depending on the source, between 42% and 58% of participants had never participated in any social movement before (Monterde, 2015) and only around 15-20% had previously participated in a political organization (Monterde, 2015; Anduiza et al., 2014).

2. They all employed social networking sites, combined with older web applications and Internet tools (Della Porta and Mattoni, 2015), without which they could not have coordinated those movements (Rheingold, 2002).

3. The online tools were used in conjunction with face-to-face gatherings and the development of radical, contentious performances, among them the physical occupation of public spaces (Della Porta and Mattoni, 2015).

2.2.2 Chronology

The movement started when the organization “Democracia Real YA!” (Real Democracy NOW!) called for a demonstration on May 15, 2011, with the slogan “We are not merchandise for politicians and bankers. Real Democracy NOW!”. After the creation of the Facebook event, a lot of other organizations spontaneously joined the call. The call had higher success than expected and eventually more than 50 Spanish cities organized a demonstration on that date. The hashtag #15mani was, on Twitter, the first trending topic (i.e. the most talked-about keyword) on May 15 in Spain, and one of the first three worldwide.

In Madrid, at the end of the demonstration, dozens of people decided to sit in Callao Street. There were confrontations with the police and 18 people were arrested. That night, between thirty and forty people, unrelated to the organizers, decided to camp in Puerta del Sol, one of the major squares of the Spanish capital. During the night between May 16 and May 17, the police evicted the camp. In solidarity with Madrid, Barcelona and other cities organized a camp that same night. The afternoon of the next day—May 17, 2011—about 10,000 people joined the camp at Puerta del Sol.

On the morning of May 19, 2011, the Junta Electoral de Madrid (Electoral Committee of Madrid) forbade demonstrations because the municipal elections were to be held on May

22. The camp in Madrid disobeyed the prohibition and stayed at the square. However, the police finally decided not to intervene. On May 20, at night, camps were spread all over Spain. In the morning of May 20, 166 camps were counted and this number reached 480 the night of this same day (Monterde, 2015). All 52 province capitals had a camp in their main square.

Once established, camps organized themselves into thematic commissions to discuss the different subjects of interest and elaborate proposals that were then discussed at the General Assembly held each day in the afternoon. During the second week of camping (May 23 to May 30, 2011), participants started to set mid-term goals: coordination of the movement at the Spanish level and its expansion at a worldwide level. After discussion,[9] two demonstrations were programmed: on June 19 and on October 15, 2011. The former was the most important and most attended event of the movement. Participants demonstrated against the Pact for the Euro[10] and demanded a European Union for the citizens and not for the markets. The October 15 demonstration was designed to expand the movement beyond Spain and the call to demonstration was also made at the international level.

At the end of May, participants decided to start decentralizing the movement away from the main squares of the cities. By the end of the week, the movement started its articulation in the neighborhoods and municipalities of each town, where "local councils" were held. The camps in squares were lifted voluntarily by the participants between the end of May and the beginning of August depending on the city, stating "we are not leaving, we are expanding".

2.3 Political context

Since the mid-eighties, the Spanish government has been held in alternation by two political parties: PSOE (left-wing party) and PP (right-wing), with around 35% of votes each. The far-left party IU (with communist origins) obtained around 6%, the Catalan nationalist parties around 4% and the Basque nationalists parties around 1.5% (see Figure 1).

After 2015, the clear hegemony of these two parties was broken by the emergence of new or previously marginal parties that obtained a notable proportion of votes. One particularly notable example of this is Podemos—literally *"We can"*— that became the third political force in Spain with around 20% of the vote.

Podemos was created in March 2014 by a group of political scientists from Madrid. It became the third largest party within the first 20 days it allowed membership, with

[9] At first, it was proposed to organize a big march that would start from every large town to converge on July 17 to "take over Madrid and re-found democracy". Finally, this initiative did not prosper.

[10] A new political general framework for the implementation of structural reforms intended to improve competitiveness, employment, financial stability and the fiscal strength of each country in the EU that required structural reforms of pensions, toughening fiscal rules and transposing and implementing one of the Stability and Growth Pact's fiscal rules directly into national legislation to make it more effective.

100,000 signing up in that period. The first election for which the party ran were the European elections of 2014. It is further to the left compared to PSOE and seeks to address the problems of inequality, unemployment and economic malaise that followed in the wake of the European debt crisis. Representatives of the party refer to 15M in their discourse and claim that without 15M, Podemos could never have existed.[11] During the 2014 European elections, Partido X, another less promiment party also closely related to the *Indignados* movement, also ran for for the first time.[12]

In 2019, Vox, another notable political actor, gained a strong electoral base. Vox is a far-right political party created in 2013. It remained marginal until 2017 when it gained importance in the media scene after the Eurosceptic summit where the party leader participated along other far-right leaders like Marine Le Pen (FN) or Frauke Petry (AfD) (Ferreira, 2019). The April 2019 congressional elections ended the Spanish exception of being one of the only European countries not having a far-right party present in the Parliament, with Vox obtaining around 10% of the vote.

3 Data

Data was collected from several sources to construct a unique city-level dataset of protests, weather, demographic variables, and electoral results in Spain between 2004 and 2019. Table E1 shows basic descriptive statistics for the main variables used in the analysis.

Electoral Results The electoral results at the city level for both the congressional and the European elections are collected from official sources published on the website of the Interior Ministry of Spain for the period from 2004 to 2019.[13] The vote share of each political option is computed as the proportion of the people in the electoral census (i.e. all people that had the right to vote) that choose to vote for that option in each city. I aggregate electoral options along two axes: left to right and anti-corruption to not anti-corruption, as described in Section 4 and Appendix D. I focus on the electoral results of the Congress and the European elections for several reasons. First, they are the two electoral events that are the most homogeneous at the national level, with campaigns focusing on nation-wide issues and similar parties running in all electoral districts. Second, it is the Congress that has the most relevant legislative functions and it elects the president of the Spanish government. Third, the European elections of 2014 were the very first

[11]See El Mundo "Sin este movimento [el 15M], Podemos no hubiese sido posible" or Ideal "Podemos somos herederos del 15M"

[12]Partido X is a small party directly issued from 15M (see for exemple El País, "El partido X se quita la máscara"; El confidencial: "Falciani ficha por el Partido X, que sale del "anonimato" la semana que viene").

[13]The Spanish legislative system has two chambers: a Congress and a Senate. In general elections, Spaniards choose a total of 350 deputies (congressmen and congresswomen) that will elect the president of the Spanish government. They also elect 208 of the 259 senators.

election for which parties claiming to be heirs of 15M (Podemos and Partido X), and closely associated to 15M by participants (Table E2) were an electoral option.

Protest measures I use as my treatment variable the number of participants in the June 19th march in each Spanish city that hosted one. I obtained the number of people in each demonstration from three different sources: 1) from the online website and Facebook page of "Democracia Real YA" (Real Democracy NOW), a group that was very close to the organizers of the rally; 2) from a compilation available online on the webpage "Toma la Plaza" (Take the Square) which is also related to the organizers; 3) from my own compilation of different local and national press sources. The final database contains 89 demonstrations (see Figure 2). I describe in Appendix D.1 how I combine these sources to obtain a single measure of attendance.

Weather data The weather data is extracted from the online meteorological service World Weather Online that collects historical weather measured every 3 hours. I extracted the quantity of rain (in mm) and the apparent temperature (i.e. the equivalent temperature felt by a human being, accounting for wind and humidity) every 3 hours in all the cities of the sample. I create a dummy variable for unpleasant weather equal to one if the day of the demonstration was either hot or rainy. I classify a day as hot if the apparent temperature was greater or equal to 30ºC at 12pm or 3pm. I classify a day as rainy if at least 0.1mm of precipitation has been observed at 12pm or 3pm. I also use the weather data to create three weather controls: i) the probability of observing unpleasant weather in an average day of June, estimated from the weather in June from 2010 to , excluding 2011; ii) the weather conditions in the same month of the marches, estimated using the probability of observing unpleasant weather in all days of June 2011 three days before and three days after the day of the marches and iii) the weather conditions in the months surrounding the months of the marches, estimated using the probability of observing unpleasant weather in May or July 2011. I detail the reasons behind the choice of the unpleasant weather measure in Appendix B.

Twitter data Twitter data was collected using the Twitter Academic Research API. To examine different dimension of the social network activity, I collected multiple samples based on different search criteria. To observe general social network activity after June 19, 2011, I collected a random sample of one million tweets by collecting tweets containing the 100 most common words in Spanish during random intervals in the month following the protest.[14] I also collected during the same interval the set of all tweets containing words and hashtags related to 15M movement and the June 19 demonstration.[15] As a

[14] The full list of words is shown in Table D1.

[15] The list of hashtags has been built by initially searching for #15M and expanding the set by looking for the most frequently associated hashtags, keeping only hashtags that are not tied to a particular geographic area. The final list is presented in Table D2.

placebo, I also collected a random sample in April 2011, before the start of the movement. To study the later social network activity around the movement, I also collected tweets containing "15M" or "Podemos", either as a word or a hashtag and including retweets, during the months preceding the 2015 congressional election help on December 20, and the month preceding the April 28, 2019 election. Finally, I collected all tweets containing the Spanish words for "corrupt" and "corruption" during the week following major corruption scandals before and after the movement. For each tweet, I extract the time and text of the tweet, the user, the user's stated location, and account creation date. I assign tweets to a geographical location based on the location stated by the user in their profile. Not all users state a location and among those who do, not all state a valid location (e.g. "in the heart of Justin Bieber") so I restrict the sample to the users that state a valid location that can be matched to a Spanish city. The location is an arbitrary text field which is not meant to be machine-readable. I use the Nominatim geocoding engine (based on the Open Street Map database) to find the coordinates of the most likely match for the location. I then filter out all locations outside Spain and all locations that are too imprecise (e.g. "Spain"). I map the remaining coordinates to cities using geographical information from the Spanish National Geographical Institute.

Facebook network To study the online network around 15M and its proximity to Podemos, I use data on the network of likes between Facebook pages close to the 15M movement, the different political parties and the main unions. I use the Facebook Graph API to extract the pages that a particular page likes and the pages that those pages like up to level 2 (i.e. one page between the main page and the last page). To estimate the proximity in the Facebook network between the movement and each particular party, I first identify the 8 principal Facebook pages most linked to the movement and the official Facebook page of each political party. For each page linked to 15M and each political party page, I compute the ratio of the number of pages liked by both the 15M page and the party page over the number of pages liked by the party. Then, for each party, I compute the mean of those ratios for all 8 pages linked to 15M.

Pre-existing political preferences To control for pre-existing political leanings of the voters, I use data on voting intentions just prior to the movement. The data comes from a quarterly survey organized by Spain's Center for Sociological Research (CIS). Participants are asked which party they would vote for if the general elections took place the following day. Since the city of respondents is not systematically available, I aggregate this information at the province level. To have enough data for each province, I averaged the results in the four surveys preceding the 15M demonstrations (July 2010, October 2010, January 2011 and April 2011).

Corruption concern and political self-classification I use data on public concern about corruption coming from the CIS monthly survey conducted on around 2500 adult Spaniards. Participants are asked about the three most important problems of Spain. I use the same CIS survey to collect information on respondents self-classification in a political left-right 1-10 scale. Again, the exact city of interviewees is not systematically available so I aggregate this information at the province level.

Electoral surveys To gather more information about the motivation of voters, I exploit the pre and post-electoral surveys conducted by the CIS before and after each general election on a larger sample. Participants are localized at the province level. In particular, I exploit information about their past votes, their voting intentions, their evaluation of different political leaders, and whether they have an account on several social networking platforms.

Major corruption cases I collected a list of the most important corruption scandals in Spain in the 2008-2016 period from an article from El Pais.[16] I use the 7 most serious cases as assessed by the panel of experts consulted by El Pais (cases Gurtel, Punica, Palau, Barcenas, Pujol, Pretoria and ERE).

15M participants behavior To provide context for the movement and support my hypotheses and the interpretation of the results, I use data about the behaviors and beliefs of the participants of the movement from an online survey conducted in 2014 (Redes, Movimientos y Tecnopolítica, 2014).[17]

Other data I obtain data on the population of each city from the population census conducted in Spain in 2010-2011. Data on unemployment rate is not available at the city level, so I use the ratio of unemployed people over the population aged 16 to 65 as a proxy for the city unemployment rate. The GDP for each province is obtained from the National Institute of Statistics. This data is not available at the city level either, so I use it at the province level. I also obtain bilateral internal migration counts at the province and semester level from Spain's CIS.

4 Empirical Strategy

To study the effect of attendance at marches on electoral choices, I would ideally like to estimate the following regression for different political options:

$$\text{Vote share}_{ce} = \gamma_0 + \gamma_1 \frac{\text{Protesters}_c}{\text{Population}_c} + \mathbf{X}_c' \beta + \eta_e + \varepsilon_{ce}. \quad (1)$$

[16]"Cuáles son los casos de corrupción más graves de España", , *El Pais*, June 29, 2017

[17]The microdata is available under Creative Commons license on the web: Tecnopolitica (in Spanish)

Vote share$_{ce}$ is the vote share for the political option being considered, in city c, during election e. I aggregate all political parties along two axes: left *vs* right and anti-corruption *vs* non-anti-corruption. The left-right axis is divided into four categories: left, center, right and far-right.[18] The anti-corruption axis is divided into two categories: parties that explicitly mention in their electoral program the fight against corruption and those who do not. Protesters$_c$ is the number of participants in the demonstration of June 19, 2011 in city c. Population$_c$ is the total population of the city. $\mathbf{X}_c$ are city-level controls measured just before the demonstration tool place including: population, previous political behavior and unemployment levels of the city c. η_e are election fixed effects. Finally, ε_{ce} is the error term.

The coefficient of interest is γ_1. A simple OLS estimation of this equation will be biased because of endogeneity. For example, regions with a greater tradition of social mobilization are likely to behave differently in elections compared to regions with a lower tradition of social mobilization. To overcome this problem, I instrument the number of demonstrators by weather shocks at the city level during the day of the June 19 marches. The excluded instrument is unpleasant weather (either rainy or too hot). The idea behind using unpleasant weather as an instrumental variable is that when faced with unpleasant weather conditions people tend to do fewer outdoors activities (Bélanger et al., 2009; Tucker and Gilliland, 2007), including protesting (Zhang, 2016; Madestam et al., 2013) and participating in political activities (Artés, 2014).

The growing literature that makes use of weather as an IV raises concerns about violation of the exclusion restriction assumption (Mellon (2021), Gallen (2020)). The main argument is that the wide variety of variables which were previously linked to weather in other studies may be causally associated with one's dependent variable. Since these variables are affected by weather, this violates the exclusion restriction. In the following subsection I will explain all the elements and tests that make the exclusion restriction unlikely to be violated in our case. I detail the explanation further in Appendix A.

4.1 Weather as an instrument

The main specification of the first stage is as follows:

$$\begin{aligned}\frac{\text{Protesters}_c}{\text{Population}_c} = \alpha_0 &+ \alpha_1 \text{Pleasant weather}_c \\ &+ \alpha_2 \text{Probability of unpleasant weather}_c \\ &+ \alpha_3 \text{Probability of unpleasant weather}_c^2 \\ &+ \alpha_4 \text{Probability of unpleasant weather in June 2011}_c \\ &+ \mathbf{X}_c' \alpha_5 + \varepsilon_c\end{aligned} \tag{2}$$

[18]Far-right is included in the "right" category. I include it in the analysis because the rise of far-right parties is a widespread phenomena in Europe.

where Unpleasant weather$_c$ is a dummy variable equal to one if the weather was unpleasant (either rainy or with a perceived temperature higher than 30ºC) during the march in city c; Probability of unpleasant weather$_c$ is the number of days with unpleasant weather over total number of days in June during the period 2010-2014 (excluding 2011); Probability of unpleasant weather^2_c is the squared term of the previous term and accounts for non-linearity in the effect of weather conditions; Probability of unpleasant weather in June 2011$_c$ is the number of days with unpleasant weather over total number of days in June 2011, excluding three days before and three days after the exact day of the marches; all other variables are the same as in equation (1). I expect a negative sign for α_1. All standard error are robust to heteroscedasticity.

The first stage results, with different variations of the controls, are presented in Table 1. The inclusion of the probability of unpleasant weather as a control is particularly important. Indeed cities that experience unpleasant weather more frequently (for example, inner regions of the South of Spain), are likely to have different protest behavior and different electoral results than cities with lower probability of unpleasant weather (for example, areas in the Atlantic coast). The inclusion of the probability of unpleasant weather means the instrument can be interpreted as a weather shock, i.e. the deviation from the probability of experiencing an unpleasant weather on June 19, 2011 with respect to an average month of June in the city. The inclusion of a squared term produces a flexible functional form able to capture non linear effects of weather conditions and increase precision. As shown in column 2 and 4 of Table 1 and in Appendix A, results are similar I when don't including the quadratic term but the F statistic becomes smaller, creating concerns about the instrument being weak.

The inclusion of the probability of unpleasant weather on June 2011 captures a different phenomenon than the probability of unpleasant weather in other years. As weather is temporally correlated, experiencing deviation from mean weather conditions on the day of the demonstration is likely correlated to weather deviation in the same direction in the surrounding days and weeks. The exclusion restriction could be violated if weather conditions in surrounding days affect other characteristics that can impact political preferences. Controlling for these weather shocks alleviates this concern and get us closer to being able to interpret the instrument as capturing only short-run effects of weather. I exclude the six days (three days before and three days after) surrounding the day of the demonstration to leave enough variation on weather conditions. As shown in columns 4 and 5 of Table 1, this control is significant in explaining the attendance to demonstration. However, its inclusion does not significantly change the first stage or second stage results (as discussed in Appendix A). I discuss further the possible concerns of serially correlated weather and the steps I do to alleviate them in the paragraphs that follow.

Importantly, our instrument focuses on short-run weather conditions (I only include weather condition during few hours on the day of the demonstration) instead of looking at weather during a long period of time. If weather were not serially correlated, this would ensure that only causal channels that depend precisely on the weather on June 19 can lead to an exclusion violation. This would already automatically exclude any concern about the direct effect of weather on elections (Gomez et al., 2007; Artés, 2014; Persson et al., 2014) because the closest elections we consider are more than five month apart from marches.

However, even not considering temporal correlation, the presence of other major events on the same day that are impacted by weather (for example other marches) could violate the exclusion restriction. I am not aware of any other major events that happened on this day. In the same line, unpleasant weather conditions could be correlated to activities that increase awareness of the marches (such as watching the news or being on social media, activities that are mostly done indoors). In this case the violation of the exclusion restriction will lead to a downward bias: we will be comparing cities with unpleasant weather (that we would consider less exposed to the 15M movement but that in fact would be more exposed than thought) with cities with pleasant weather (that we consider as more exposed to 15M). The results of this paper are likely to be underestimated in different ways; if the example above is true, that will add another source of downward bias but will not challenge the existence of the results we find.

Weather on the same day could also influence the media coverage of the demonstration: if weather was pleasant, more journalists could have been present in the demonstration, leading to increased attention independently of any variation in the number of demonstrators. The direction from this bias could go either way: media could be more willing to cover demonstrations where the weather was pleasant, but people protesting despite unpleasant weather might be a more interesting story. However, it is still unlikely that this had an effect on the coverage of the demonstrations: the 15M movement was a well-covered topic in part due to the existence of the camps for more than a month and the demonstration was planned far in advance. Indeed, the June 19 protest was a major news story and was featured prominently on the cover of local and national newspapers of all political leanings the next day.[19]

In reality, weather patterns *are* temporally correlated. Even if we consider weather shocks, it is possible that summer 2011 was a deviation from the mean weather in Spain. If, for example, the summer was particularly hot, the exclusion restriction could be violated through other channels. For example, it could provoke a general productivity drop that leads to a decrease in income.

In addition to including the probability of unpleasant weather on June 2011 to alleviate exclusion restriction concerns related to the correlation of weather over time, as a robustness check, I confirm that results also hold when controlling for weather in the months around June 2011. In particular, I control for weather in May and July. This rules out a big part of the concerns related to the fact that weather is correlated with variables such as income, long-term pollution, mental health problems or migration that

[19]For instance in El País, La Vanguardia, El Mundo, La Razón and Sur.

can affect electoral results. In addition to controlling for weather around the time of the marches, I conduct a sensitivity analysis (following Mellon (2021)) to confirm that other causal channels detected in the literature are unlikely to be strong enough to drive the results (see Appendix A).

Events that are affected by weather that occurred in the three days around marches can violate the exclusion restriction. The 15M movement organized permanent camps, so it could be that good weather encouraged people to visit the camps in the days surrounding the marches. In that case the exclusion restriction would be violated because the good weather would affect electoral behavior not just through marches but also through camp visits and participation. This violation of the exclusion restriction would lead to an upward bias. Under the assumption that the effect of the march goes in the same direction as the effect of attendance at camps, the estimates would capture not just the effect of the march but the effect of the camps as well. I argue that this is not a big problem because even if true, I will still be capturing an effect driven by the 15M movement even if through other channels that uniquely through marches.

Finally, it is also possible that unexpected unpleasant weather can be, by chance, spuriously correlated with other characteristics affecting protest behavior and electoral choices at the same time. In the robustness check section, I present a set of results that help rule out this possibility and further validate the identification assumption.

The fifth column of Table 1 reports results for the main first stage used for the rest of the empirical study. As expected, the estimates show a statistically significant negative relation between unpleasant weather during the demonstration of June 19, 2011, and the number of participants in this demonstration. The F statistic of the first stage is higher than the critical value given by Stock and Yogo (2005) (i.e. 10) under which we should be concerned about weak instruments. However, I still conduct robustness checks using errors robust to weak instruments (Appendix A). The magnitude of the estimates is quite high: experiencing rain or a temperature higher than 30ºC is associated with participation decreasing by 2.1% of the population of the municipality. The estimated numbers of demonstrators are presented on Figure 3. In an average city, unpleasant weather reduces protest attendance by around 50%. This result is consistent with media coverage of the demonstration that reports strong concerns among participants due to weather conditions or even the need of medical intervention for some protesters due to health problems related to high temperature. [20]

[20] The media emphasized the intense heat during the demonstration: "The marches [...] took place under an intense heat" (El confidencial, 19/06/2011); "[...] heat is the main element to fight [...] participants ask for all types of drinks and buckets of ice through various social networks [...]" (RTVE, 19/06/2011); "[...] the high temperatures [...] made it necessary for the emergency services to intervene for anxiety attacks in several participants in the march" (20minutos, 19/06/2011).

4.2 Robustness and choice of instrument

In Appendix A, I conduct several robustness checks to increase confidence in the identification strategy. I first start by discussing in more details the exclusion restriction for the instrument. Besides a theoretical discussion (following Cinelli and Hazlett (2020a)), I include additional weather controls for the months surrounding the marches to control for long term weather shocks that can be linked to different outcomes themselves related to electoral results such as income, pollution or crime (Mellon, 2021). I also conduct a sensitivity analysis to ensure that even in the case there were other, unaccounted for, variables related to weather on June 19 that are linked to electoral results five months later, it is very unlikely that those links could overturn the results. Because unemployment was particularly high in Spain at the time of the marches and it can both affect protest participation (Kern et al., 2015) and electoral results (Rosenstone, 1982), I explicitly test that weather shocks on the day of the march are not linked to unemployment in our particular case (Stein, 1990; Nam, 2007; Burden and Wichowsky, 2014). I also examine whether removing some of the weather controls affects the second stage results.

In addition, I check that weather shocks on other days cannot predict the level of attendance at the marches. Indeed, if we draw the distribution of the effects of weather shocks of 240 placebo dates on the level of march attendance, we see that the distribution is centered at zero and that the effect of weather shocks on the day of the march is outside the 95% of the observations. This further assures that weather on the day of the march is, on average, not correlated to weather shocks on other days that could affect other variables also linked to electoral results.

In Appendix B, I detail the reasons for the choice of instrument, including the choice of thresholds and the weather control.

4.3 Second stage

The specification of the second stage is as follows:

$$\begin{aligned}\text{Vote share}_{ce} = \beta_0 + \beta_1 \frac{\widehat{\text{Protesters}_c}}{\text{Population}_c} \\ &+ \beta_2 \text{Probability of unpleasant weather}_c \\ &+ \beta_3 \text{Probability of unpleasant weather}_c^2 \\ &+ \beta_4 \text{Probability of unpleasant weather in June 2011}_c \\ &+ \mathbf{X}_c' \beta_5 + \eta_e + \varepsilon_{ce}.\end{aligned} \tag{3}$$

where all the variables are as explained above in equation (1) and (2). β_1 is our coefficient of interest. I compute the IV estimates using 2SLS estimates, All standard errors are computed clustering at the level of the treatment (i.e. at the city level) and are robust to heteroscedasticity.

5 Results

5.1 Main results: electoral choices

Table 2 shows the estimated impact of the rate of attendance in cities that hosted a march on June 19, 2011 on the results of all subsequent congressional and European elections up to 2019.[21] Political options are aggregated along two axes: left vs right (columns 1-4) and anti-corruption vs not anti-corruption (columns 5-6). The different panels show different election year sub-samples. Panel A shows the results for all elections (European and to the Spanish Congress) from 2011 to 2019. Panel B shows the results for all elections from 2014 to 2019. Panel C shows the results pooling the 2019 elections together (two election to the Congress and one European election). The analysis of these different sub-samples allows to capture and distinguish four important things: i) the overall effect, with the advantage of having more precise estimates due to a higher number of observations (panel A); ii) the effect once the political offer includes new political parties (not present in 2011) that are particularly linked with the Indignados movement (Podemos and Partido X) (Panel B); iii) the effect after the rise of the presence of far-right parties in the media and in the public debate (Panel C); and iv) the effects in the last available year where elections occurred, to test the long term persistence of the results (Panel C).[22]

Results are consistent throughout the different sub-samples. Having a higher number of protesters with respect to the population increases the share of vote for left-wing and anti-corruption options and decreases the vote for the far-right. An increase of one percentage point in the proportion of the city population that attends the march leads, on average, in each subsequent election, to an increase of one percentage point in the vote share for left-wing parties, and half a percentage point for anti-corruption parties. It also decreases the vote for far-right parties by an average of around a third of a percentage point. Alternatively, an increase of one standard deviation in the number of protesters raises the vote share by around 3.3 pp for left-wing parties and 2.8 pp for anti-corruption-parties, and decreases by around 0.9 pp the vote share for the far right. In absolute terms, 10 additional demonstrators lead to around 8 additional votes for left-wing parties, 6 additional votes for anti-corruption parties, and 2 votes less for far-right parties in each subsequent election, on average.[23]

[21] OLS estimates and reduced form regressions are presented in Table E3. Scatter plots of the second stage are presented in Figure E1.

[22] Appendix C shows and analyses the results separating every election and the main specific political parties.

[23] These magnitudes can also be interpreted in terms of persuasion (DellaVigna and Gentzkow, 2010); namely, the effect on the percentage of the voters that would not have voted for a particular party in the absence of the Indignados movement that are convinced to vote for that party as a consequence of the movement, i.e. that are persuaded. This interpretation is detailed in Appendix C.1.

The effect is similar for the overall sample and for the sample that only includes the elections where the political options that are particularly linked to the movement are present. This does not mean that the irruption of new political parties didn't have an effect, just that the overall leftward shift is not changed. Remarkably, the effect does not vanish over time, and the estimates for the 2019 elections are not statistically different from those using the whole sample, showing long-term effects.

Overall, results suggest that the June 19 marches increased voting for anti-corruption parties and caused a leftwards shift in voting preferences across the political spectrum (more votes for the left and fewer for the far-right), and that these effects do not fade over time and do not only depend on the presence of specific electoral options. The negative effect for the far right is unexpected. There exists a several number of reason why one could think that the 15M would not have a negative effect on the far-right. First, the far-right, particularly in Spain, is very far from the demands of the 15M.In this context, voters that could be convinced away from Vox are very unlikely to be receptive to the 15M movement's message and won't be affected by the movement. On the contrary, voting for the far-right has been linked with new ways to express a form of frustration with the political system (Rooduijn et al., 2016). In this setting, we could expect information shared by 15M sympathizers to fuel this frustration, driving more voters towards Vox rather than away.[24] Finally, the right-wing vote has been linked to a higher preference for public order. In this case again, we could expect people with right-wing political preferences to shift rightwards to restore public order after the 15M street protest, increasing rather than decreasing the vote for the far-right. All these hypotheses contradict the findings of this paper. Rather, I argue that right-wing voters are convinced away from voting for far-right options, showing an effect on different parts of the political spectrum.

I use post-electoral surveys to estimate the effect of higher attendance at protests in the respondent's region on the probability of shifting the vote towards far-right parties in 2019. Results are presented in Table E4 and show that higher march attendance decreases the amount of voters that shift from voting for the right to voting for the far-right. Among voters of right-wing parties in 2016, those that lived in regions with higher protest attendance are less likely to shift their vote to the far right in 2019 elections. Unfortunately, this analysis is only possible for shifts between 2016 and 2019 elections due to the prior absence of a credible far-right option, and the unavailability of electoral surveys asking for enough electoral history.

[24]Prominent slogans from the 15M movement include "Do not vote for them" (refereeing to traditional political parties); or "Our dreams do not fit in your ballot boxes" confirming a negative sentiment towards the traditional political system.

5.2 Alternative outcomes: political placement and corruption concerns

In the previous section, I considered electoral choices as the dependent variable. In this section I use survey data to conduct a similar analysis at the individual level, focusing on slightly different but related outcomes: i) the prevalence of corruption concerns among the general public, measured as the percentage of people in a given province who consider corruption as the main problem of Spain; and ii) the political self-identification along the left-right axis (on a scale from 1 to 10). The use of survey data allows me to control for demographic factors such as age, level of education, etc.

In particular, using the same IV strategy I used in previous sections, I estimate the following equation as a second stage:

$$\begin{aligned}\text{Outcome}_{ipm} = \phi_0 + \phi_1 & \frac{\text{Protesters}_{cp}}{\text{Population}_{cp}} \\ & + \mathbf{X}'_{cp}\phi_2 \\ & + \mathbf{W}'_{ipm}\phi_3 \\ & + \lambda_m + \varepsilon_{ipm},\end{aligned} \tag{4}$$

where Outcome_{ipm} is a dummy variable capturing either i) whether individual i living in province p interviewed in month m thinks that corruption is the main problem of Spain; ii) whether individual i living in province p interviewed in month m places themselves at the left (respectively right; far-right) politically; or iii) whether individual i living in province p interviewed in month m answers "I don't know" to the question "on a 1-10 left-right scale, where do you place yourself politically?". Because the city of interviewees is not available, I am only able to estimate this regression at the level of the province instead of the city. To do that, I use the number of participants to the demonstration in the capital city of the province as a treatment. Indices p and cp refer to the province and capital of the province, respectively. $\mathbf{X}_{cp}$ and $\mathbf{W}_{ipm}$ are vectors of city- and individual-level controls respectively, and λ_m are fixed effects for the month of the interview. All other coefficients are the same as in equation 1.

Table 3 presents the results.[25] As in the main table described in the section above, different panels show different year sub-samples. Results show that higher levels of attendance in the protests cause a persistent increase in corruption concerns (column 1). In particular, an increase of 1 pp in the proportion of protesters among the population leads to an increase of around 0.42 pp in the probability of being highly concerned about corruption. Magnitudes remain very similar in the different sub-samples, indicating a persistent effect of protest attendance.

[25]The corresponding first stages are presented in Table E5, and the OLS and reduced form estimates are shown in Table E6.

In cities with higher instrumented attendance to the demonstrations, the percentage of people who self-identify as left-leaning is higher and remains higher for all year sub-samples (column 2). Magnitudes remain very similar across the samples and are not statistically different from one another (i.e. confidence intervals overlap). On average, an increase of 1 pp in the proportion of protesters in the population increases by around 1.6 pp the probability of responding that one is left-leaning. Higher attendance also decreases the probability of self-placement at the far-right. This decrease also remains constant across sub-populations, with similar magnitudes. On average, an increase of 1 pp in the ratio of protesters in a city decreases by around 0.27 pp in the probability of a person placing themselves at the far-right of the spectrum. Finally, we also observe a decrease in "does not know" answers. Magnitudes are similar to those for "left-leaning" answers, with opposite signs. For this last effect, estimates diminish over time and become marginally insignificant (but still negative) in 2019. Overall, results at the individual level are consistent with regional electoral results and confirm a persistent effect on political preferences, which are shifted towards the left, and on corruption concern, which is higher.[26]

5.3 Heterogeneity

5.3.1 Regional heterogeneity: previous political leanings

Different cities with different political preferences might react differently to protest. One could think, for instance, that right-leaning cities would react negatively to left-leaning demonstrations and vote more for the right to impose order in the public space (Huet-Vaughn, 2013), while left-leaning cities would be more likely to be convinced by the protests and vote accordingly. If the effect in left-leaning cities is larger or if there is a larger number of left-leaning cities, the overall effect could show an increase for left-wing vote but hide a polarizing effect. What are then the differential effects of higher protest attendance in cities with different political preferences?

To investigate this heterogeneity, I interact the presence of pleasant weather during 15M marches in a city with two measures of previous political preferences: i) the vote shares for left and right.wing political options in the closest national election conducted before the 15M and ii) the share of people intending to vote for the traditional left-wing and right-wing parties during the year preceding the marches[27] [28].

Table 4 shows the interaction with the vote share for left-wing (right-wing) parties in 2009 elections in Panel B1 (Panel C1) and with intention to vote for the left (right) in the year preceding the marches in Panel B2 (Panel C2). Each panel uses the full sample of

[26] Anecdotal evidence from a survey (Redes, Movimientos y Tecnopolítica, 2014) among participants in the movement goes in the same direction documenting that around a third of participants said that the Indignados movement made them change what they consider desirable or unacceptable in a society.

[27] I perform the interaction with the reduced form instead of using a 2SLS IV strategy due to concerns regarding the weakness of the instrument for the interaction terms.

[28] Intention to vote variables come from survey data. For this exercise, the share of survey respondents planning to vote for each political side is computed for each province. Analysis is done at the city level.

elections in 2011 and after. Panel A shows the reduced form for comparison. Overall both left- and right-leaning cities moved leftwards but in a different manner. Both types of cities seem to differentially increase the vote share for left-wing parties (column 1 panels B1 and C2). In addition, right-leaning cities decreased their vote for right and far-right parties (column 3 panel C1 and column 4 panels C1 and C2).

On the contrary, results for anti-corruption parties come mainly from left-leaning cities (column 5 panels B1 and B2). The results for the two different measures of past political preferences at the city level are different for several of the effects. For the great majority, they go in the same direction and show same order of magnitudes. There are two exceptions. First the decrease in the vote for the right in right-wing cities (column 3 panels C1 and C2), which have very different magnitudes even if the have the same sign. The second exception is the decease on abstention for right-leaning cities (column 7 panel C2). This could have various interpretations. However, given that this effect is just marginally significant and that it is the only significant effect that changes signs between the two different measures of previous political preferences, I interpret it as an effect of chance.

5.3.2 Individual heterogeneity: age and higher education

The population inside each region is heterogeneous. People with different characteristics choose to vote for different electoral options and react differently to protest. For example younger people tend to have lower turnout and to change political attitudes more easily as a consequence of protest (Uppal and LaRochelle-Côté, 2012; Mazumder, 2019). Restricting the analysis to the regional level can hide important effects of different signs that will be averaged out in regional results. This is especially true for minorities because even strong effects can be diluted into an indifferent majority. Yet minorities can have an important role in institutional change, for example if they are a pivotal agent or if, well organized, they can convince other groups of their views (Madestam et al., 2013; Casanueva-Artís et al., 2021).

The use of survey data allows to study individual heterogeneity based on differences between demographic groups. In particular, I study the differential effect of higher attendance on younger people and people with higher education. As in the regional heterogeneity analysis, I use the reduced form and interact living in a region that had pleasant weather during the day of the march with i) being 20 or younger during the marches and ii) having completed some form of higher education. The outcomes of interest are, again, dummy variables capturing: i) whether the respondent considers corruption as the main problem of Spain and ii) whether the respondent considers themselves as being in each of four categories (left, right, far-right, or "I don't know") on the left-right political scale.

Results are presented in Table 5 and show that people who were young when marches happened have a differential positive effect on self-identification with the left wing, and a negative effect on self-identification with the right wing (panel B). The same differential effects are seen for people with higher education (panel C). Comparing with panel A, which shows the main reduced form for reference, we see that this heterogeneity analysis revealed an effects not seen before: a significant decrease in self-identification with the right. Additionally, even if young people have a stronger effect of protest on political preferences, they have a lower effect on the level of corruption concern. I interpret this as an artifact of the specific measure of corruption concern. At the time of marches, young adults were particularly affected by a deep economic crisis (youth unemployment rates were close to 40% and it was very difficult four young people to quit the parental house due to precarious work conditions and high housing prices). Corruption concern is a dummy variable that measures whether a person thinks that corruption in the *main* problem of Spain, not just that it is *a* problem. It is likely that youths had other major concerns besides corruption even if they also worry about corruption.

Overall, the heterogeneity analysis confirms the interpretation that higher levels of attendance to the protests led to a leftwards shift of the whole distribution of political preferences. Categories of people who would have had high probabilities of being left-wingers even without the protests (namely, people living in regions that were more left-leaning just before the 15M movement, younger people, and people with higher education) shifted further to the left on average (voting more for the left and less for the right, and identifying themselves more with the left and less with the right). On the other hand, right-leaning categories (people living in regions with higher intention to vote for the right before the protests) voted less for the far-right.

5.4 Placement in the literature

Overall, the results support the hypothesis that the main demonstration of 15M changed voting behavior according to the movement's demands and claims. The sign of the effect is consistent with other findings in the literature, however the magnitudes are smaller. Madestam et al. (2013) find, for example, that one additional demonstrator in the rallies organized by the Tea Party on 2009's Tax day added 12 votes for the Republican Party in the midterm elections held one and a half years after the rallies. Larreboure and González (2021) find that one additional protester at the Women's March added 3 votes for women and minorities candidates in the midterm elections of 2018, almost two years after the march. I find smaller effects. I find that one additional protester lead to around 1 additional vote for left-wing parties, 0.4 additional votes for anti-corruption parties, and 0.3 fewer votes for far-right parties in all subsequent elections up to eight years after the demonstration. Part of this difference in magnitude is likely due to the different time frame that I consider (up to eight years instead of less than two for the other two papers) but a part of it could also be due to an underestimation of the effect in this paper. Unlike the other two papers, this paper measures the effect of having a larger number of demonstrators with respect to having fewer demonstrators but not with respect to not having a demonstration at all. More generally, my results could be underestimating the effect of 15M. Indeed, I only consider the effect of one day of demonstrations, and I compare relative outcomes across different cities, rather than measuring the impact of protests on Spain as a whole.

5.5 OLS vs 2SLS

In all estimation results, we see that the magnitudes of the OLS (shown in Table E3) are lower than for the IV. This could be due to either a downward bias of the OLS or to the difference between the Average Treatment Effect (ATE) estimated by the OLS and the Local Average Treatment Effect (LATE) estimated by the IV (Imbens and Angrist, 1994). Intuitively, one may think that the OLS would be biased upwards instead of downwards as the characteristics of a city (for example: region with a history of high past protest activity or a higher number of activist associations) may be correlated with the electoral behavior of their citizens (in our example: more likely to vote for left-wing options) and with their probability of having a high number of demonstrators (more likely to protest). Thus, in this case, I argue that the differences between OLS and 2SLS results come from the difference between ATE and LATE. The LATE gives the effect on the compliers instead of the average effect on the population. Compliers are defined as the individuals who change their behavior after a treatment. In this case the treatment is the exposure to a weather shock and the behavior we are interested in is the attendance at a march. Thus, in our case the compliers are the individuals that would have passed over the march if there was a negative weather shock and attended the march because there was not a negative weather shock or vice-versa. These individuals are more likely to experience a stronger effect of the demonstration precisely because they are the ones that were less convinced to protest to begin with. Let us take two extreme examples. Person A is a far-left activist that attended the march, that would had attended the march regardless of weather conditions and that would have voted for the left regardless of the march. The effect of the march on this person is zero. Consider now person B, who places at the center/center-left of the political spectrum. This person is not completely convinced by attending the march and they would had preferred to stay at home if it was a rainy day. It turns out that the weather was particularly nice on the day of the march and person B attended the march. Normally, they would vote for the center but they were very convinced by the arguments they heard during the march and they changed their electoral choice to a left-wing option. The effect of attending the march for this person is different from zero. While the ATE give us the average effect for the population (the average between person A and person B), the LATE give us the effect just for person B. Since the effect for B is positive and the effect for A is zero, in this case the OLS will give us a lower estimate than the 2SLS using weather shocks as an instrument. This is a simplified example at the individual level. Our data is at the city level, but the argument for individuals could be translated to cities: we capture the effect of the share of compliers in the city.

5.6 Robustness of the main results

I conduct several robustness check for the main results in Appendix A. I start by conducting a placebo test showing that the level of attendance at marches cannot predict the electoral results in the elections that happened before the march. I deal with possible concerns about outliers by conducting the analysis excluding one city at a time and showing that results remain virtually the same in every case. I additional add non-core controls (GDP and dummy variables for two regions that consistently behave differently in elections); I consider errors robust to weak instruments, account for possible spatial correlation and correct standard errors for multiple hypothesis testing.

6 Mechanisms: new social media networks

As with most of the Occupy movements, 15M participants demonstrated an up-to-date knowledge about socialization dynamics on the Internet and the technical aspects that underlie virtual platforms (Martínez Roldán, 2011; Della Porta and Mattoni, 2015). The importance of communication was well understood, and there was a clear will among participants to self-mediatize the protest and its demands (Castells, 2015). In the most important camps during the heyday of the movement, communication commissions were organized and put in charge of communicating about the movement's activities and its main claims[29]. These commissions developed strategies to reach Twitter's "trending topics". The call to the June 19 march, for example, was preceded by new daily hashtags for a week, each referring to different issues related to the movement but always with the date of the demonstration. This systematic communication discipline made the discussion around the demonstration accessible to a large number of users, reaching out to new communities and spreading the call effectively. The movement became the most popular trending topic on Twitter for several days during the heyday of the protest.[30] The camps of the different cities created Facebook pages that achieved a large number of followers within weeks.[31] Data coming from a survey conducted to march participants (Redes, Movimientos y Tecnopolítica, 2014) documents that around a third of interviewed participants started to use Twitter due to the 15M movement (Table E7).

In this section, I argue that the 15M street protests had long term political effect

[29] In Madrid, for example, at least two persons were working constantly to communicate information about the movement and its achievements.

[30] El movimiento 15-M, del anonimato al 'trending topic'

[31] acampadaBCN (Barcelona) reached 70,000 followers and acampadaSol (Madrid camp) 44,000 followers in the first two months of the movement.

because they were successful in creating an online network that keeps exposing people to a content close to the 15M opinions.

I will proceed in several steps. First I show that cities with higher attendance at the marches created more Twitter accounts in the month following the demonstration. Then I show that those cities tweeted more about 15M. Third, I show that this regional difference in activity persists over time. In particular people from cities with higher protest attendance still tweet more about 15M; tweet more about corruption after corruption scandals; and tweet more about the main left-leaning party linked to the 15M movement (Podemos) before elections.[32] Finally, I turn to the survey data and show that the effect of protest attendance on electoral choices are stronger for people with social media accounts.

6.1 Higher social media activity uptake and differential activity

I hypothesize that higher attendance at protests encourages social media uptake. I test this hypothesis using Twitter data. I collect and geo-localize at the city level a random sample of tweets containing the 100 most common Spanish words.[33] Each tweet contains information about the author and the creation date of the account, which I exploit to compute, for each city, the number of distinct users and the number of distinct accounts created after the demonstration on June 19th.

I aggregated all of these different measures into an index using the first Principal Component Analysis (PCA).[34] I regress the PCA index as well as the three components separately on the level of protest attendance using the same IV strategy. Results are presented in Table 6. Panel A shows a placebo analysis: it captures the effect on the same variables measured in April 2011, before the 15M movement started. None of the measures are significant, showing that cities with a higher and lower protest attendance were not different in terms of Twitter activity before the 15M movement. Panel B presents the variables measured during the month following the demonstration and shows that higher protest attendance increases the index of the level of Twitter activity in the month that follows (column 1). Protest attendance increases activity at the intensive margin: the overall number of tweets – talking about any topic but containing one of the 100 most Spanish used words– (column 2) increases, as well as the number of distinct active users, tweeting about any topic– (column 3). Higher protest attendance also increases Twitter activity at the extensive margin. More new accounts were created in cities with more protesters (column 4).

Is this increase related an increase in 15M content? To answer this question, I collected and geo-localized all tweets containing 15M related hashtags in the month following the marches. Table 7 shows the effect of the number of protesters on the level of Twitter activity around the 15M during the following month. Estimates show that higher levels

[32] See section 2.3 for a brief description of the party and its creation.
[33] The complete list of words can be found in Appendix Table D1.
[34] The description of the factors identified by PCA is shown in Table D3.

of protest attendance lead to more 15M-related tweets (column 1) and a higher number of distinct users tweeting about 15M (column 2). Column 3 shows that this is not only an increase in negative activity: tweets were classified between positive, negative and neutral using sentiment analysis (see §D.3), and column 3 shows that the number of tweets expressing a positive sentiment also increases.

6.2 Persistence in 15M related social network activity

15M-related discourse I test the persistence of regional variation in Twitter activity in Table 8. In panel A, I consider the activity during the month preceding the 2015 congressional election, and in panel B, the month preceding the 2019 congressional election. Columns 1 and 2 show that more tweets about 15M were sent by a larger population in the period preceding both the 2015 and the 2019 elections in cities with higher protest attendance. Magnitudes decreased between the period following the demonstrations and 2015 but remain similar between 2015 and 2019.

Left-wing party linked to 15M: Podemos I then turn to the analysis of the activity around the political party most linked to the 15M: Podemos. Again, I collect and geo-localize all tweets containing hashtags related to Podemos before the 2015 and 2019 elections. I show that regions with higher protest attendance tweet more about Podemos: the number of tweets is higher (column 3) and the number of distinct users as well (column 4).

6.2.1 Corruption and corruption scandals

I now focus on the public debate around corruption. I test whether people talked more or are more concerned about corruption in cities more affected by marches. To do so, I collect and geo-localize all tweets mentioning corruption during one week following each of a set of major corruption scandals reported by the media, hypothesizing that even if protests increased concern about corruption in general, people will particularly express those concerns after a *trigger*. I consider four major corruption scandals after the 15M movement, with the last one occurring in 2018, seven years after the movement.I regress the logarithm of the number of tweets about corruption and distinct users talking about corruption on the level of protest attendance. Results are presented in Table 9. Panel A presents the results for the number of tweets while panel B presents the estimates for the number of distinct users. For all corruption scandals, estimates are positive and significant showing that users in cities exposed to larger protest are talking more about corruption in the online arena even in 2018, almost seven years after the demonstrations.

I then analyze corruption concerns in the offline scene. To do that I use survey data asking about the level of concern about corruption. Figure 4 shows the evolution of the percentage of people that think that corruption is the first problem that affects Spain in regions with pleasant and unpleasant weather during the June 19 march. There is a clear difference between before and after the movement. In all semesters after the movement, people are more concerned about corruption than before the movement. Before the movement, we do not observe any difference between regions that had pleasant and unpleasant weather during the June 19 march, but after the movement, regions with pleasant weather (thus bigger demonstrations) are more concerned about corruption than regions with unpleasant weather. Moreover, even when the concern for corruption diminishes nationally, there is no convergence between regions with pleasant and unpleasant weather on the corruption concern in the eight years after the movement.

After each of the two major corruption cases that were revealed after 15M, the difference in corruption concerns between regions with pleasant and unpleasant weather becomes larger, even long after the movement. That the protesters themselves have become more sensitive to this issue during the protests does not seem sufficient to explain this increased difference long after 15M. The existence of social network links that were created during the protests does.

Going beyond a simple graphical analysis, I estimate the effect of having larger demonstrations on corruption concerns in different periods. I replicate the analysis done in equation (3) by splitting the data in six months periods. Figure 5 shows the estimates of equation (3) with 90% confidence intervals for six month periods before and after the movement. Estimates are close to zero before the movement and most are statistically insignificant. In contrast, after the movement, all are positive and 10 out of 16 are statistically significant. This means that regions with larger marches are more concerned about corruption than regions with lower attendance, even more than four years after. We also observe that the effect is larger in periods with corruption scandals and does not disappear even after eight years.

For the six-month periods following each corruption scandal, I show the (individual level) regression results in Panel C of Table 9. Again, the results are positive and significant. The results from the regressions in Panel A and Panel C are represented on Figure 6, along with estimates of the effect for scandals that occurred before the 15M movement: for these placebo results, the effects of demonstrations are not statistically significant.

This change in people's main concerns is another (indirect) way in which the existence of social networks can influence the results of elections. Indeed, previous literature has documented an important role of information transmission on electoral behavior (DellaVigna and Kaplan, 2007), including information about corruption (Peters and Welch, 1980).

6.3 Social media and electoral results

Interaction In the previous subsections, I have established a link between higher protest attendance and a higher use of social media, specially about 15M related topics. However, observing an increase in Twitter activity related to 15M does not necessarily mean than this activity caused a change in political preferences. I use the post-electoral surveys from Spain's Centro de Investigaciones Sociológica to better understand the role of social media as a channel explaining effects of protests. Along with their vote in the latest elections, survey participants are asked whether they have an account on several social media platforms. I use the reduced form to interact pleasant weather with having a social media account and examine the differential effect of being a social media user on the vote for different political options.[35].

Results are presented in Table 10 and show that having a social media account can increase the effect of pleasant weather during the marches on electoral results. In particular, people with a social media account vote relatively more for the left and for anti-corruption options and relatively less for centrist or right-wing parties due to pleasant weather conditions in their city during the demonstration. Results are persistent at least up to 2019 when considering the effect on voting for the center and the left. I hypothesize that the lack of long-term differential effects for some of the options are due in part to the reduced number of observations and in part due to a reduction in variation in the people with and without a social media account.[36] Indeed, as time passes, more and more people join social media so the *"control"* group composed by people with no social media account reduces and changes in composition.

These results cannot be interpreted causally: having a social media account might be related to other unobservable factors making individuals more sensitive to the 15M's message, even when controlling by individual characteristic such as age, gender, education level and labor situation. However, they indicate that voters having access to social media have different voting patterns that align more strongly with the 15M goals if the marches were larger, which is compatible with them being exposed to the 15M's message through its social media presence.

Facebook network Until now I have used only Twitter data to establish the uptake of social media, I will here look at the network on Facebook. Even if I show in previous subsections that regions with higher protest attendance tweet more about 15M and more about Podemos, it is possible that the 15M community and the Podemos community are parallel in the sense that they coexist but do not interact. In this subsection, I use both Twitter *and* Facebook data to argue that the two communities are linked.

I first show that before the 2015 elections (the first Congressional elections in which Podemos runs for), the ratio of tweets mentioning 15M among tweets mentioning Podemos is higher in cities with higher protest attendance. The results are shown in column 1 of Table E8. This means that people are associating more the new party Podemos with 15M in those cities.

[35] I use the reduced form due to concerns with weak instruments in the interactions.

[36] For right-wing political options, the coefficient for 2019 remains virtually the same as in the other panels but standard errors increase, possibly due to a reduction in the number of observations.

Second, I use Facebook data to gather the network of *"likes"* around 15M and around selected political parties. In particular, I collected the list of all Facebook pages that liked several 15M-related Facebook pages and the list of all pages that those pages liked (so I collected the data at the second level).[37] The same process is used for pages related to Podemos and other prominent political parties less closely related to 15M. At the end, I obtain, for each entity (political party or 15M), a set of pages that are in their respective "network". Finally, I look at the ratio of likes of 15M that overlap with each political party. Results are shown in Figure 7 and show that the 15M network is far more linked with the Podemos network than with the network of any other party.

7 Alternative mechanisms

7.1 Internal migration based on political preferences

Protests can be signals of political preferences of a city (Lohmann, 1993) and voters can "vote with their feet" and decide to move to cities with political preferences closer to theirs (Tiebout, 1956; Hirschman, 1970). In the context of 15M protests, voters can migrate across cities. For example, left-leaning voters could observe the success of the protests in another city and prefer moving to this city instead of moving to a city with lower protest attendance. In that case, part of the long-term effects could come from an increase in left-wing voters in cities with higher number of protesters. Right-wing voters may similarly decide to move to regions with lower protest attendance, or regions where voters are more aligned with their preferences (e.g. being repelled by some policies decided at the municipality level). We cannot directly observe the political preferences of internal migrants. However, if voters decided to move because their city's politics don't align with their views, this migration would be more important between cities where the protests were differently successful. In the reduced form, this would show up as a move from a city with pleasant weather to a city with unpleasant weather, or from a city with unpleasant weather to a city with pleasant weather on the day of the demonstration.

I exploit province-to-province bilateral internal migration data to argue that internal migration does not drive our results. Due to limitations of the available data, I consider migration from province to province, and use the weather in the main demonstration of a province. I use the following specification to see whether migration flows are more important between cities with different weather:

$$Y_{od} = \alpha_1 \text{Different weather}_{o,d} + \alpha_2 \text{Distance}_{o,d} + \alpha_3 \text{Distance}^2_{o,d} + \omega_o + \delta_d + \varepsilon_{o,d}$$

where Y_{od} is a measure of migration flows from origin province o to destination province

[37] Unfortunately, the Facebook API does not allow collecting of the list of pages that like one page anymore. That is why this analysis is done using data from the networks from 2015.

d and Different weather$_{o,d}$ is a dummy equal to one if province o had pleasant weather and province d unpleasant weather on the 19th of June, or conversely if o had unpleasant weather and d pleasant weather. Distance$_{o,d}$ is the distance between the two provinces' capitals, ω_o are origin-province fixed effects and δ_d are destination-province fixed effects. I control by distance, because while characteristics of the origin or destination province are absorbed by fixed effects, characteristics of pairs of provinces still need to be accounted for: migration between distant provinces is less likely than migration between neighboring provinces. The standard errors are computed by two-way clustering along both the origin province and the destination province (Cameron et al., 2011).

The results are presented in Table E9 for different measures of migration between the second half of 2011 and the end of 2021. In column 1, I use the simple count of migrants from origin to destination. To take into account the size of the province, I divide in column 2 by the population of the origin province, and in column 3 by the number of emigrants from the origin province. The likelihood of migrating to a given destination province also depends *a priori* on how large the destination province is: column 4 takes the value of column 2, and additionally divides by the share of the Spanish population represented by the destination province, thus accounting for the size of the destination province. Similarly, column 5 takes the value of column 3 and divides it by the share of internal migrants going to the destination province, thus taking into account the attractivity of the destination province. In all cases, the coefficient of having a different weather is not significantly different from zero.

Instead of voters choosing to move out of their province because its political preferences doesn't match theirs and going to another province closer to their preferences, it is possible that voters that had already chosen to move for unrelated reasons may choose different destinations based on their political preferences (Batut and Schneider-Strawczynski, 2022). If both left-wing people migrate preferentially to left-wing regions, and right-wing people migrate preferentially to right-wing regions, the net effect on bilateral internal migration flows may be null, as both changes may cancel each other. To observe such a phenomenon, we would need to observe the composition of the migration flows instead of only observing counts.

7.2 Inertia

There are other mechanisms that could explain in part or in full the long-term effects of protest. Electoral inertia is one such mechanism: good results for a party in one election may causally lead to good results in subsequent elections.

This could be due to effects on both the offer and demand side of the electoral market. On the offer side, good electoral results can lead to a set of advantages in the next election including higher campaign budgets and better-known candidates (Erikson, 1971). Moreover, candidates that are successful in a location might invest more campaigning effort there in the next election, or even advantage this location more in their policy choices, expecting higher returns (Curto-Grau et al., 2012). On the demand side, multiple mechanisms may also be at play. For example, voters may be reluctant to reevaluate their past votes (Mullainathan and Washington, 2009; Meredith et al., 2009). If they are voting strategically, they will consider previously-successful candidates as more viable options. It is also possible that part of the long-term effect is due to sticky habit formation. The literature has already documented the creation of sticky political preferences after specific outside events that can be influenced by weather conditions such as the celebration of the 4th of July in the United States (Madestam and Yanagizawa-Drott, 2012) or past elections (Fujiwara et al., 2016).

I use individual level survey data to try to explore the inertia mechanism further. In particular, I take advantage of post-electoral surveys that ask respondents who they voted for in the past two elections. This allows me to restrict the sample to voters of a certain party and explore who they voted for in the next election. I take as a particular example Podemos voters. Table E10 presents the results. Columns 1 restricts the sample to Podemos voters in a certain election and shows the effect of unpleasant weather on the probability of switching away from voting Podemos (i.e. voters that voted Podemos in a certain election and did not vote for Podemos in the next election any more). Column 2 restricts the sample to voters that do not vote Podemos and shows the effect on the probability of switching from any other electoral option to Podemos (i.e. voters that voted for any political party except Podemos and that voted for Podemos in the subsequent election). Each panel corresponds to a different parliamentary election. Panel A present the results for 2016 elections depending on the vote in 2015. Panel B presents the results for the April 2019 election depending on 2016 elections and panel C presents the results for November 2019 elections depending on the vote in April 2019. Results show that pleasant weather during the day of the demonstration (and thus a higher relative number of demonstrators) increases the switch to voting Podemos from other political options in April 2019 with respect to 2016 (column 2, panel B) and in November 2019 with respect to April 2019 (column 2, panel C). These results suggest that, at least on the demand side, the full effect of unpleasant weather does not simply come from an initial increase in the vote for Podemos that sticks. Instead, this shows that weather conditions during the day of the demonstrations continue to have an effect in switching voters to Podemos eight years after the marches.

It is important to note that this result at the individual level is not able to capture the offer side of the mechanism that can be at play (as discussed at the beginning of this subsection). These results are only able to give empirical evidence suggesting that electoral inertia may not explain in full the existence of long-term effects of the relative attendance at marches on the electoral behavior of voters.

7.3 Other possible mechanisms

Beyond the two mechanisms already considered, there may be other channels that can explain the effects. For example, the messages of the 15M movement could spread not only through social networks, but also through offline social connections, for example among friends and family. One can imagine an extreme example where an offline social network is created around the 15M march. This new social network is very active offline but, additionally, the members of the network also created links online. It is however, their offline activity that leads to changes in political preferences. In this case, one would find results consistent with the online network analysis mechanism even if the online network does not mediate the results and is just a good proxy for the off-line network. In the same line but beyond the extreme example that I present above, protests might also act as a recruitment channel for forms of political activism that are not limited to social media participation. As more people attend the demonstrations, more potential activists may be recruited, which leads to local voters being more exposed to the idea of 15M, and subsequently of Podemos if these activists move to more formal political participation after the creation of Podemos.

Protests can act as an information shock and may inform about local political preferences. Higher attendance at the demonstration could be interpreted as a sign of a higher prevalence of a certain political preference. This information shock can lead to a change in the behavior of different political agents. First, it may affect the offer side (i.e. the political parties). For example, Podemos could have focused their campaign efforts on regions where the 15M demonstrations have been larger, trying to capture as much as possible the vote of 15M sympathizers. Second, it may have affected the demand side (i.e. the voters). For example, strategic voters can start considering Podemos, which is related to 15M, as a more viable electoral option, especially in the first elections where Podemos was running. Finally, this informational shock may affect local media coverage. By observing attendance at the demonstration, editors might conclude that local audiences are interested in news about 15M, and are concerned about corruption. This could lead to a change in coverage in an attempt to better match this audience (Zhuang, 2019), and subsequently to more people being exposed to news about 15M, corruption, and eventually about the creation of Podemos.

8 Conclusion

Protests against the political system have been rising during the last decade all over the world, achieving unprecedented levels of mobilization. Existing research shows important short term effects of protest on various political outcomes. In this paper, I extend this literature by showing that modern protests can have long term effects on political preferences and by proposing, testing and documenting a new possible channel that explains part of these lasting results: the creation of an online social network as a result of street protests.

Using cross-city variation in weather shocks to instrument the level of protest participation on the day of the march, I estimate the causal effects of variation in attendance at simultaneous marches organized by the Spanish Occupy movement 15M on subsequent electoral results and political opinions up to a decade following the protests. I find that relatively larger marches shifted political preferences leftwards and towards parties that are explicitly anti-corruption while increasing public concern about corruption at least up to 2019.

To explain the channels leading to long-lasting effects, I use social media data to show that the 15M street protests were successful in creating an online network that exposes people to a content close to the opinions of the movement in a persistent way. I show that higher levels of protest attendance lead to higher Twitter uptake and to more online debate around 15M related topics. I then show that this change in the nature of the online debate was persistent. Analysis of both Twitter and survey data shows that people were more concerned about corruption and tweeted more about it in regions with higher protest attendance after each major corruption scandal while corruption concern and online discourse were not different between those regions before. Finally, I show using survey data that voters that have a social media account react more strongly to the size of the demonstration, even when controlling for individual characteristics.

Taken together, the evidence shown in this paper supports the idea that one-time non-violent protests such as large simultaneous marches can change people's electoral behavior (in this particular case towards the left and anti-corruptions options) and that this effect can persist and affect the local electoral behavior for several years, particularly if social networks are created around the protests.

References

1. **Acevedo, Sebastian, Mico Mrkaic, Natalija Novta, Evgenia Pugacheva, and Petia Topalova**, "The effects of weather shocks on economic activity: what are the channels of impact?," *Journal of Macroeconomics*, 2020, *65*, 103207.

2. **Allcott, Hunt, Luca Braghieri, Sarah Eichmeyer, and Matthew Gentzkow**, "The welfare effects of social media," *American Economic Review*, 2020, *110* (3), 629–76.

3. **Anderson, Theodore W and Herman Rubin**, "Estimation of the parameters of a single equation in a complete system of stochastic equations," *The Annals of Mathematical Statistics*, 1949, *20* (1), 46–63.

4. **Anduiza, Eva, Camilo Cristancho, and José M Sabucedo**, "Mobilization through online social networks: the political protest of the indignados in Spain," *Information, Communication & Society*, 2014, *17* (6), 750–764.

5. **Arnold, Felix and Ronny Freier**, "Only conservatives are voting in the rain: Evidence from German local and state elections," *Electoral Studies*, 2016, *41*, 216–221.

6. **Artés, Joaquín**, "The rain in Spain: Turnout and partisan voting in Spanish elections," *European Journal of Political Economy*, 2014, *34*, 126–141.

7. **Barberá, Pablo**, "How social media reduces mass political polarization. Evidence from Germany, Spain, and the US," *Job Market Paper, New York University*, 2014, *46*, 1–46.

8. **Barbieri, Francesco, Luis Espinosa Anke, and José Camacho-Collados**, "XLM-T: A Multilingual Language Model Toolkit for Twitter," *CoRR*, 2021, *abs/2104.12250.*

9. **Batut, Cyprien and Sarah Schneider-Strawczynski**, "Rival guests or defiant hosts? The local economic impact of hosting refugees," *Journal of Economic Geography*, 2022, *22* (2), 327–350.

10. **Bélanger, Mathieu, Katherine Gray-Donald, Jennifer O'loughlin, Gilles Paradis, and James Hanley**, "Influence of weather conditions and season on physical activity in adolescents," *Annals of epidemiology*, 2009, *19* (3), 180–186.

11. **Boxell, Levi, Matthew Gentzkow, and Jesse M Shapiro**, "Greater Internet use is not associated with faster growth in political polarization among US demographic groups," *Proceedings of the National Academy of Sciences*, 2017, *114* (40), 10612–10617.

12. **Brannen, Samuel, Christian Haig, and Katherine Schmidt**, "The age of mass protests: understanding an escalating global trend," 2020.

13. **Brox, Enzo and Tommy Krieger**, "Far-right protests and migration.," 2021.

14. **Buechler, Steven M**, "New social movement theories," *Sociological Quarterly*, 1995, *36* (3), 441–464.

15. **Burden, Barry C and Amber Wichowsky**, "Economic discontent as a mobilizer: unemployment and voter turnout," *The Journal of Politics*, 2014, *76* (4), 887–898.

16. **Cameron, A Colin, Jonah B Gelbach, and Douglas L Miller**, "Robust inference with multiway clustering," *Journal of Business & Economic Statistics*, 2011, *29* (2), 238–249.

17. **Campante, Filipe, Ruben Durante, and Francesco Sobbrio**, "Politics 2.0: The multifaceted effect of broadband internet on political participation," *Journal of the European Economic Association*, 2017, *16* (4), 1094–1136.

18. **Campbell, Travis**, "Black Lives Matter's Effect on Police Lethal Use-of-Force," *Available at SSRN*, 2021.

19. **Casanueva-Artís, Annalí, Vladimir Avetian, Sulin Sardoschau, and Kritika Saxena**, "Going Viral in a Pandemic: Social Media and the Broadening of the Black Lives Matter Movement," *Available at SSRN 3831819*, 2021.

20. **Castells, Manuel**, *Networks of outrage and hope: Social movements in the Internet age*, John Wiley & Sons, 2015.

21. **Cinelli, Carlos and Chad Hazlett**, "Making sense of sensitivity: Extending omitted variable bias," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 2020, *82* (1), 39–67.

22. **_____ and _____** , "An omitted variable bias framework for sensitivity analysis of instrumental variables," *Work. Pap*, 2020.

23. **_____ , Jeremy Ferwerda, and Chad Hazlett**, "Sensemakr: Sensitivity analysis tools for ols in r and stata," *Available at SSRN 3588978*, 2020.

24. **Conley, T.G.**, "GMM estimation with cross sectional dependence," *Journal of Econometrics*, 1999, *92* (1), 1 – 45.

25. **Curto-Grau, Marta, Alfonso Herranz-Loncán, and Albert Solé-Ollé**, "Pork-Barrel Politics in Semi-Democracies: The Spanish "Parliamentary Roads," 1880–1914," *The Journal of Economic History*, 2012, *72* (3), 771–796.

26. **de Amorim, Guilherme Marques**, "Communication networks and protests: investigating the "Occupy Movement" in the United States," 2016.

27. **DellaVigna, Stefano and Ethan Kaplan**, "The Fox News effect: Media bias and voting," *The Quarterly Journal of Economics*, 2007, *122* (3), 1187–1234.

28. ______ **and Matthew Gentzkow**, "Persuasion: empirical evidence," *Annu. Rev. Econ.*, 2010, *2* (1), 643–669.

29. **Dimant, Eugen and Guglielmo Tosato**, "Causes and effects of corruption: what has past decade's empirical research taught us? A survey," *Journal of economic surveys*, 2018, *32* (2), 335–356.

30. **El-Mallakh, Nelly, Mathilde Maurel, and Biagio Speciale**, "Arab spring protests and women's labor market outcomes: Evidence from the Egyptian revolution," *Journal of Comparative Economics*, 2018, *46* (2), 656–682.

31. **Enikolopov, Ruben, Alexey Makarin, and Maria Petrova**, "Social media and protest participation: Evidence from Russia," *Econometrica*, 2020, *88* (4), 1479–1514.

32. **Erikson, Robert S**, "The advantage of incumbency in congressional elections," *Polity*, 1971, *3* (3), 395–405.

33. **Ferreira, Carles**, "Vox como representante de la derecha radical en España: un estudio sobre su ideología," *Revista Española de Ciencia Política*, 2019, (51), 73.

34. **Fominaya, Cristina Flesher and Laurence Cox**, "Fighting for a voice: the Spanish 15-M/Indignados movement," in "Understanding European Movements," Routledge, 2013, pp. 256–273.

35. **Fontenla, Matías, M Ben Goodwin, and Fidel Gonzalez**, "Pollution and the choice of where to work and live within Mexico City," *Latin American Economic Review*, 2019, *28* (1), 1–17.

36. **Fujiwara, Thomas, Kyle Meng, and Tom Vogl**, "Habit formation in voting: Evidence from rainy elections," *American Economic Journal: Applied Economics*, 2016, *8* (4), 160–88.

37. **Gallen, Trevor**, "Broken instruments," *Available at SSRN 3671850*, 2020.

38. **Garz, Marcel and Gregory J Martin**, "Media Influence on Vote Choices: Unemployment News and Incumbents' Electoral Prospects," *American Journal of Political Science*, 2021, *65* (2), 278–293.

39. **Gomez, Brad T, Thomas G Hansford, and George A Krause**, "The Republicans should pray for rain: Weather, turnout, and voting in US presidential elections," *The Journal of Politics*, 2007, *69* (3), 649–663.

40. **González, Felipe**, "Collective Action in Networks: Evidence from the Chilean Student Movement," 2017.

41. **Hirschman, Albert O**, *Exit, voice, and loyalty: Responses to decline in firms, organizations, and states*, Vol. 25, Harvard university press, 1970.

42. **Huet-Vaughn, Emiliano**, "Quiet riot: The causal effect of protest violence," *Available at SSRN 2331520*, 2013.

43. **Imbens, Guido W. and Joshua D. Angrist**, "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 1994, *62* (2), 467–475.

44. **Kern, Anna, Sofie Marien, and Marc Hooghe**, "Economic crisis and levels of political participation in Europe (2002–2010): The role of resources and grievances," *West European Politics*, 2015, *38* (3), 465–490.

45. **König, Michael D., Dominic Rohner, Mathias Thoenig, and Fabrizio Zilibotti**, "Networks in Conflict: Theory and Evidence From the Great War of Africa," *Econometrica*, 2017, *85* (4), 1093–1132.

46. **Larreboure, Magdalena and Felipe González**, "The impact of the Women's March on the US House Election," *Documento de Trabajo IE-PUC*, 2021, (560).

47. **Levy, Ro'ee**, "Social media, news consumption, and polarization: Evidence from a field experiment," *American economic review*, 2021, *111* (3), 831–70.

48. **Levy, Roee and Martin Mattsson**, "The effects of social movements: Evidence from# MeToo," *Available at SSRN 3496903*, 2021.

49. **Liu, Yinhan, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov**, "Roberta: A robustly optimized bert pretraining approach," *arXiv preprint arXiv:1907.11692*, 2019.

50. **Lohmann, Susanne**, "A signaling model of informative and manipulative political action," *American Political Science Review*, 1993, *87* (2), 319–333.

51. **Lu, Xun and Halbert White**, "Robustness checks and robustness tests in applied economics," *Journal of Econometrics*, 2014, *178*, 194 – 206. Annals Issue: Misspecification Test Methods in Econometrics.

52. **Madestam, Andreas and David Hans Yanagizawa-Drott**, "Shaping the nation: The effect of Fourth of July on political preferences and behavior in the United States," *HKS Faculty Research Working Paper Series*, 2012.

53. ——— **, Daniel Shoag, Stan Veuger, and David Yanagizawa-Drott**, "Do political protests matter? Evidence from the tea party movement," *The Quarterly Journal of Economics*, 2013, *128* (4), 1633–1685.

54. **Manacorda, Marco and Andrea Tesei**, "Liberation technology: mobile phones and political mobilization in Africa," 2016.

55. **Mazumder, Soumyajit**, "The Persistent Effect of US Civil Rights Protests on Political Attitudes," *American Journal of Political Science*, 2018, *62* (4), 922–935.

56. ——— , "Black Lives Matter for Whites' Racial Prejudice: Assessing the Role of Social Movements in Shaping Racial Attitudes in the United States," 2019.

57. **Mellon, Jonathan**, "Rain, Rain, Go Away: 176 potential exclusion-restriction violations for studies using weather as an instrumental variable," *Available at SSRN 3715610*, 2021.

58. **Meredith, Marc et al.**, "Persistence in political participation," *Quarterly Journal of Political Science*, 2009, *4* (3), 187–209.

59. **Mikusheva, Anna and Brian P Poi**, "Tests and confidence sets with correct size when instruments are potentially weak," *The Stata Journal*, 2006, *6* (3), 335–347.

60. **Ministerio del Interior**, "Las elecciones generales en España 1977-2016," Technical Report Dec 2017.

61. **Monterde, Arnau**, "Emergencia, evolución y efectos del movimiento-red 15M (2011-2015). Una aproximación tecnopolítica." PhD dissertation, Universitat Oberta de Catalunya 2015.

62. **Mullainathan, Sendhil and Ebonya Washington**, "Sticking with your vote: Cognitive dissonance and political attitudes," *American Economic Journal: Applied Economics*, 2009, *1* (1), 86–111.

63. **Müller, Karsten and Carlo Schwarz**, "From hashtag to hate crime: Twitter and anti-minority sentiment," *Available at SSRN 3149103*, 2020.

64. **Nam, Taehyun**, "Rough days in democracies: Comparing protests in democracies," *European Journal of Political Research*, 2007, *46* (1), 97–120.

65. **Navarro, Vicenç**, "¿Es el gobierno Zapatero socialdemócrata?," http://www.psoe.es/izquierdasocialista/docs/546707/page/gobierno-zapatero-socialdemocrata-i.html 2011. Accessed: 31-05-2019.

66. **Persson, Mikael, Anders Sundell, and Richard Öhrvall**, "Does Election Day weather affect voter turnout? Evidence from Swedish elections," *Electoral Studies*, 2014, *33*, 335–342.

67. **Peters, John G. and Susan Welch**, "The Effects of Charges of Corruption on Voting Behavior in Congressional Elections," *The American Political Science Review*, 1980, *74* (3), 697–708.

68. **Porta, Donatella Della and Alice Mattoni**, *Spreading protest: social movements in times of crisis* 2015.

69. **Rappaport, Jordan**, "Moving to nice weather," in "Environmental amenities and regional economic development," Routledge, 2009, pp. 25–53.

70. **Redes, Movimientos y Tecnopolítica**, *#Encuesta15M2014*, Barcelona, España: IN3, Universitat Oberta de Catalunya, 2014. Recuperado de https://tecnopolitica.net/es/content/encuesta15m2014.

71. **Rheingold, Howard**, *Smart Mobs: The Next Social Revolution*, Perseus Books Group, 2002.

72. **Roldán, S Martínez**, "Movimiento 15M: construcción del espacio urbano a través de la acción de las Multitudes Inteligentes [en línea]," *Barcelona: UOC*, 2011.

73. **Romano, Joseph P and Michael Wolf**, "Stepwise multiple testing as formalized data snooping," *Econometrica*, 2005, *73* (4), 1237–1282.

74. **Rooduijn, Matthijs, Wouter Van Der Brug, and Sarah L De Lange**, "Expressing or fuelling discontent? The relationship between populist voting and political discontent," *Electoral Studies*, 2016, *43*, 32–40.

75. **Rosenstone, Steven J**, "Economic adversity and voter turnout," *American Journal of Political Science*, 1982, pp. 25–46.

76. **Rotesi, Tiziano**, "Do social media matter? The impact of Twitter on political participation," Technical Report, Mimeo 2019.

77. **Sampedro, Victor and Josep Lobera**, "The Spanish 15-M Movement: a consensual dissent?," *Journal of Spanish Cultural Studies*, 2014, *15* (1-2), 61–80.

78. **Sharoff, Serge**, "Open-source corpora: Using the net to fish for linguistic data," *International journal of corpus linguistics*, 2006, *11* (4), 435–462.

79. **Sisco, Matthew R, Silvia Pianta, Elke U Weber, and Valentina Bosetti**, "Global climate marchessharply raise attention to climate change: Analysis of climate search behavior in 46 countries," *Journal of Environmental Psychology*, 2021, p. 101596.

80. **Skoy, Evelyn**, "Black Lives Matter Protests, Fatal Police Interactions, and Crime," *Contemporary Economic Policy*, 2021, *39* (2), 280–291.

81. **Solé-Ollé, Albert and Pilar Sorribas-Navarro**, "Trust no more? On the lasting effects of corruption scandals," *European Journal of Political Economy*, 2018, *55*, 185–203.

82. **Stein, Kenneth W**, "The intifada and the 1936-39 uprising: A comparison," *Journal of Palestine Studies*, 1990, *19* (4), 64–85.

83. **Stock, James and Motohiro Yogo**, "Testing for Weak Instruments in Linear IV Regression," *Identification and Inference for Econometric Models*, 2005, pp. 80–108.

84. **Teeselink, Bouke Klein and Georgios Melios**, "Weather to Protest: The Effect of Black Lives Matter Protests on the 2020 Presidential Election," *Available at SSRN 3809877*, 2021.

85. **Tiebout, Charles M**, "A pure theory of local expenditures," *Journal of political economy*, 1956, *64* (5), 416–424.

86. **Tucker, Patricia and Jason Gilliland**, "The effect of season and weather on physical activity: a systematic review," *Public health*, 2007, *121* (12), 909–922.

87. **Uppal, Sharanjit and Sébastien LaRochelle-Côté**, *Factors associated with voting*, Statistics Canada Ottawa, 2012.

88. **Wasow, Omar**, "Agenda seeding: How 1960s black protests moved elites, public opinion and voting," *American Political Science Review*, 2020, *114* (3), 638–659.

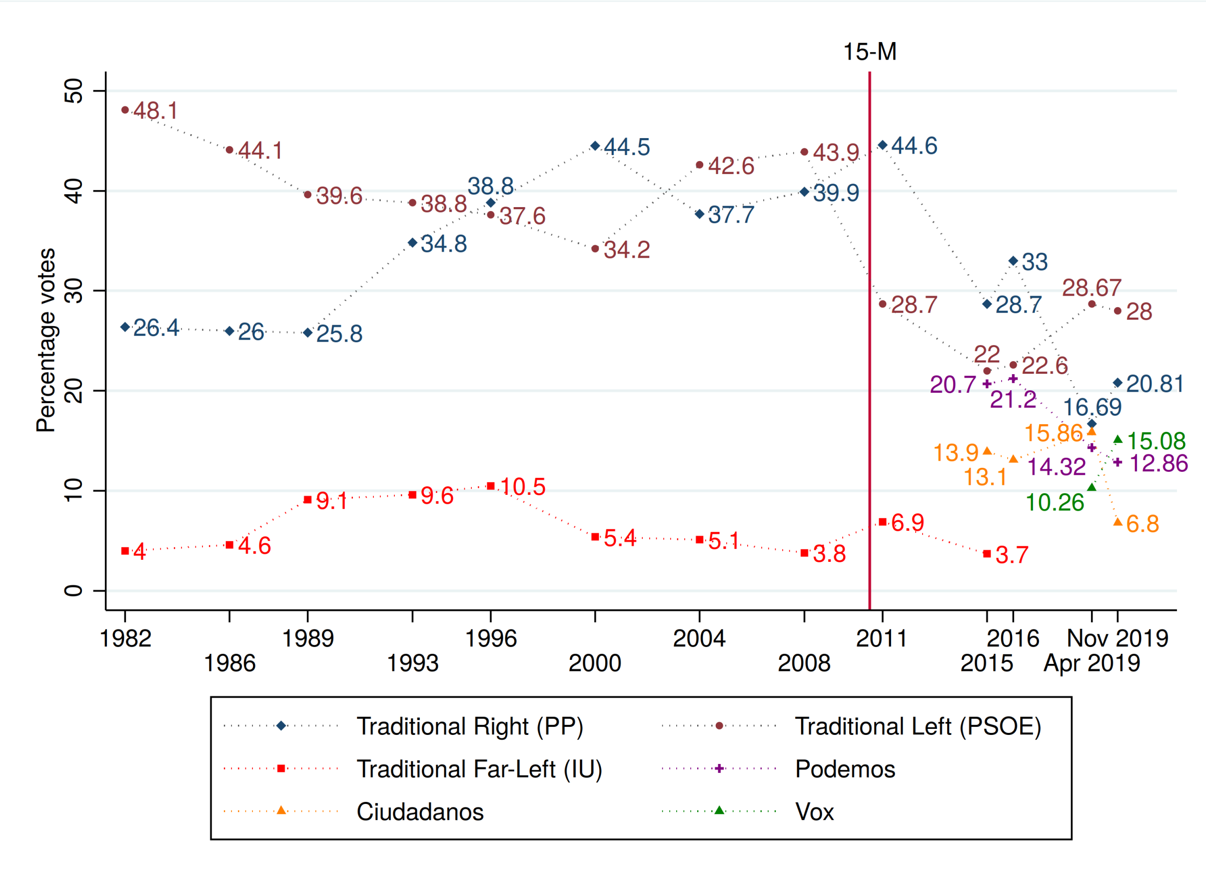
89. **Zhang, Tony Huiquan**, "Weather effects on social movements: evidence from Washington, DC, and New York City, 1960–95," *Weather, Climate, and Society*, 2016, *8* (3), 299–311.

90. **Zhuang, Maiting**, "Intergovernmental conflict and censorship: Evidence from china's anti-corruption campaign," *Available at SSRN 3267445*, 2019.

91. **Zhuravskaya, Ekaterina, Maria Petrova, and Ruben Enikolopov**, "Political Effects of the Internet and Social Media," *Forthcoming, Annual Review of Economics. DOI/10.1146/annurev-economics-081919-050239*, 2019.

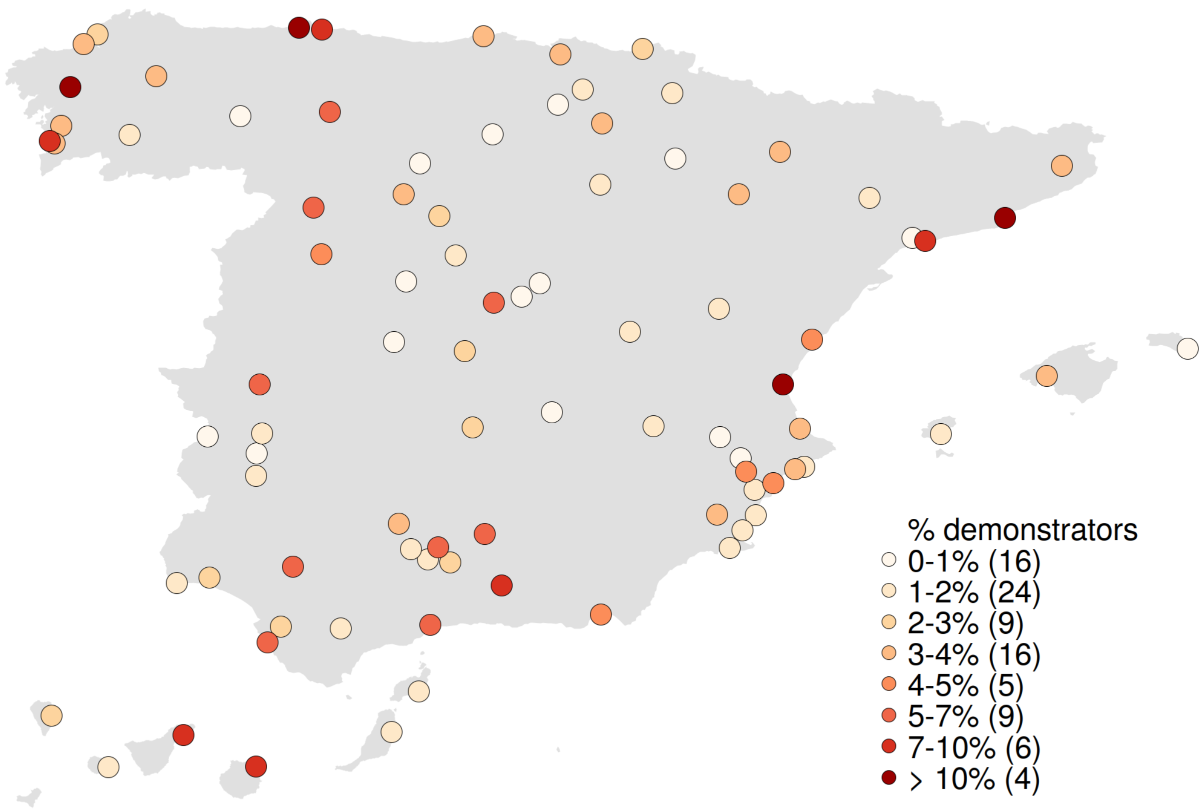
9 Figures and tables

Figure 1: Evolution of the vote shares of main political parties in Spain.



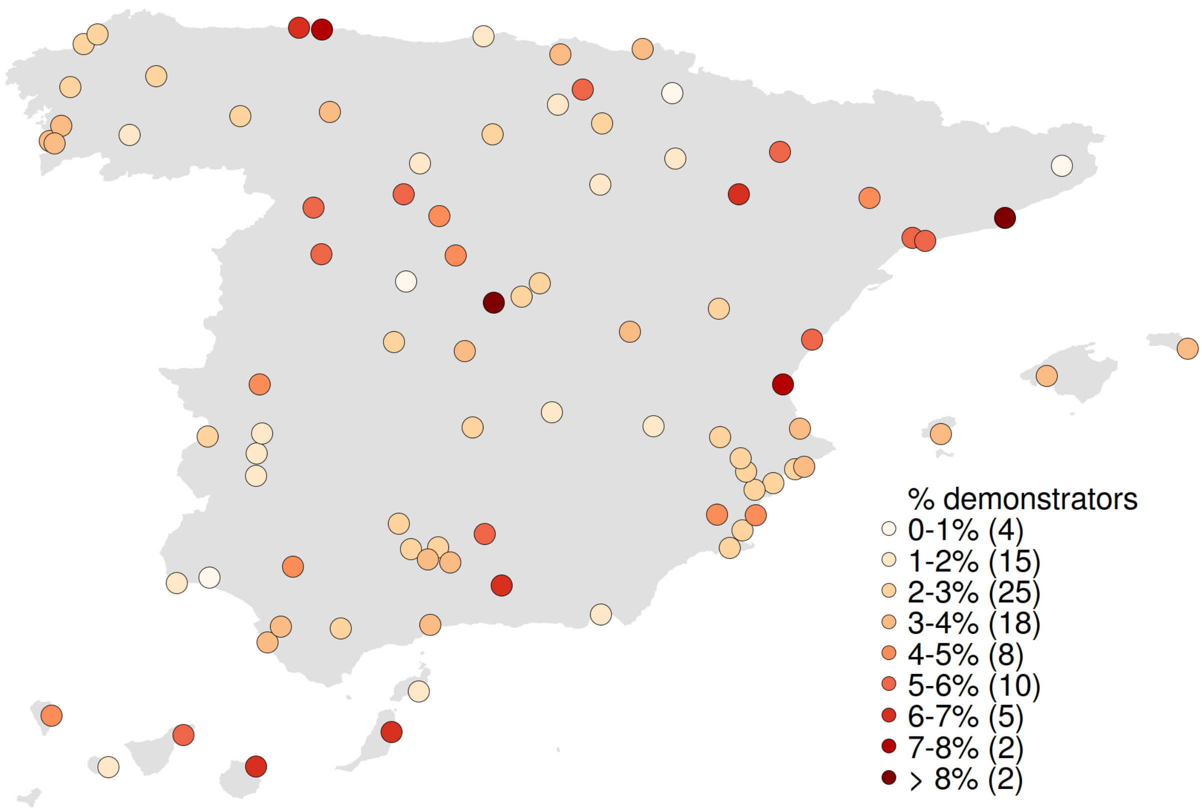
Note: Percentage of votes at the national level for major parties in Spain between 1982 and 2019. Vertical red line represents the 15M movement. Data comes from: Ministerio del Interior (2017)

Figure 2: Map of Spain: Percentage of demonstrators by city.



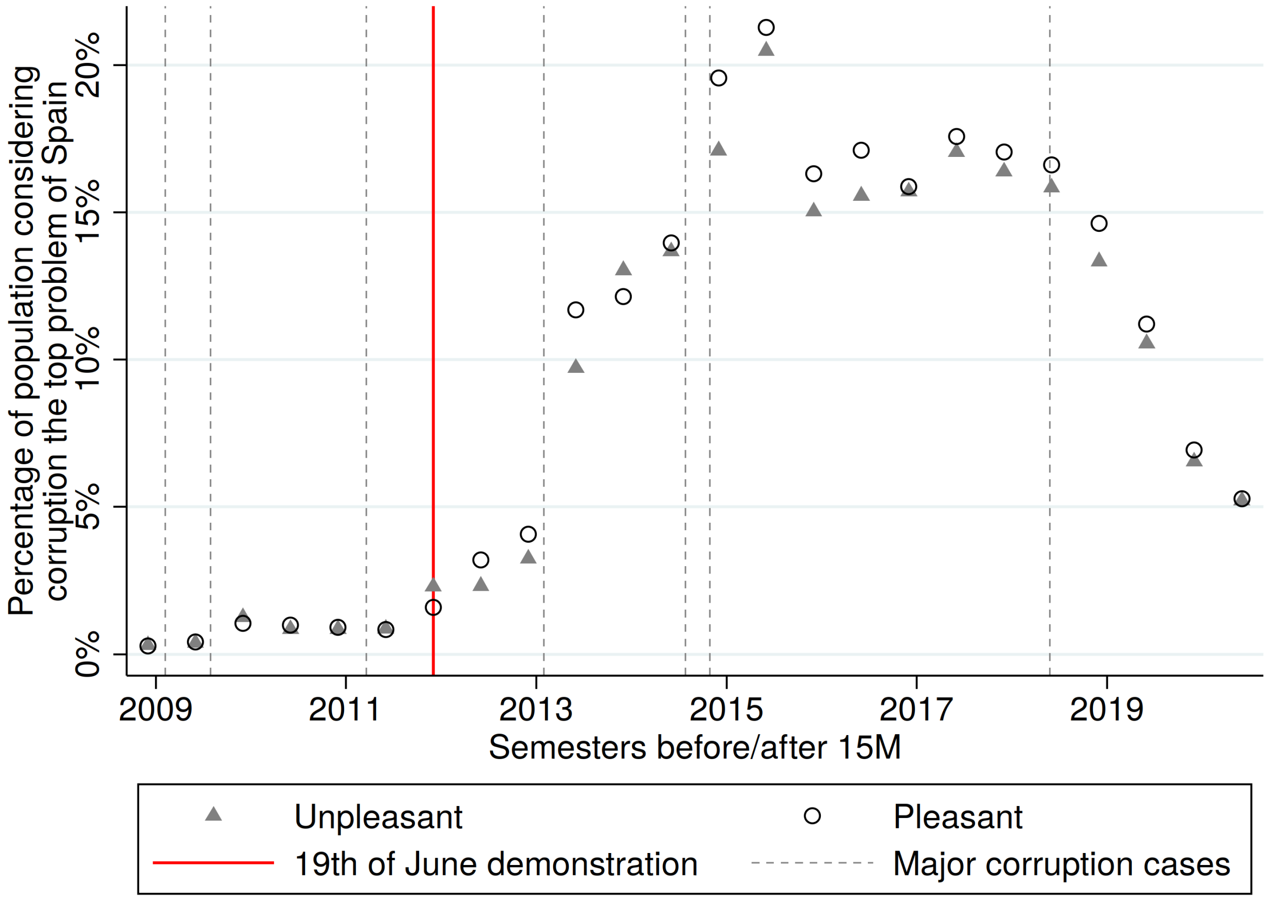
Note: Localisation of all demonstrations on June 19, 2011. Different colours represent different ratios of demonstrators to city population. Darker colours mean higher proportion of demonstrators. Number of cities for each range of values are in parentheses.

Figure 3: Map of Spain: Predicted percentage of demonstrations by city, using unpleasant weather as an exogeneous variation.



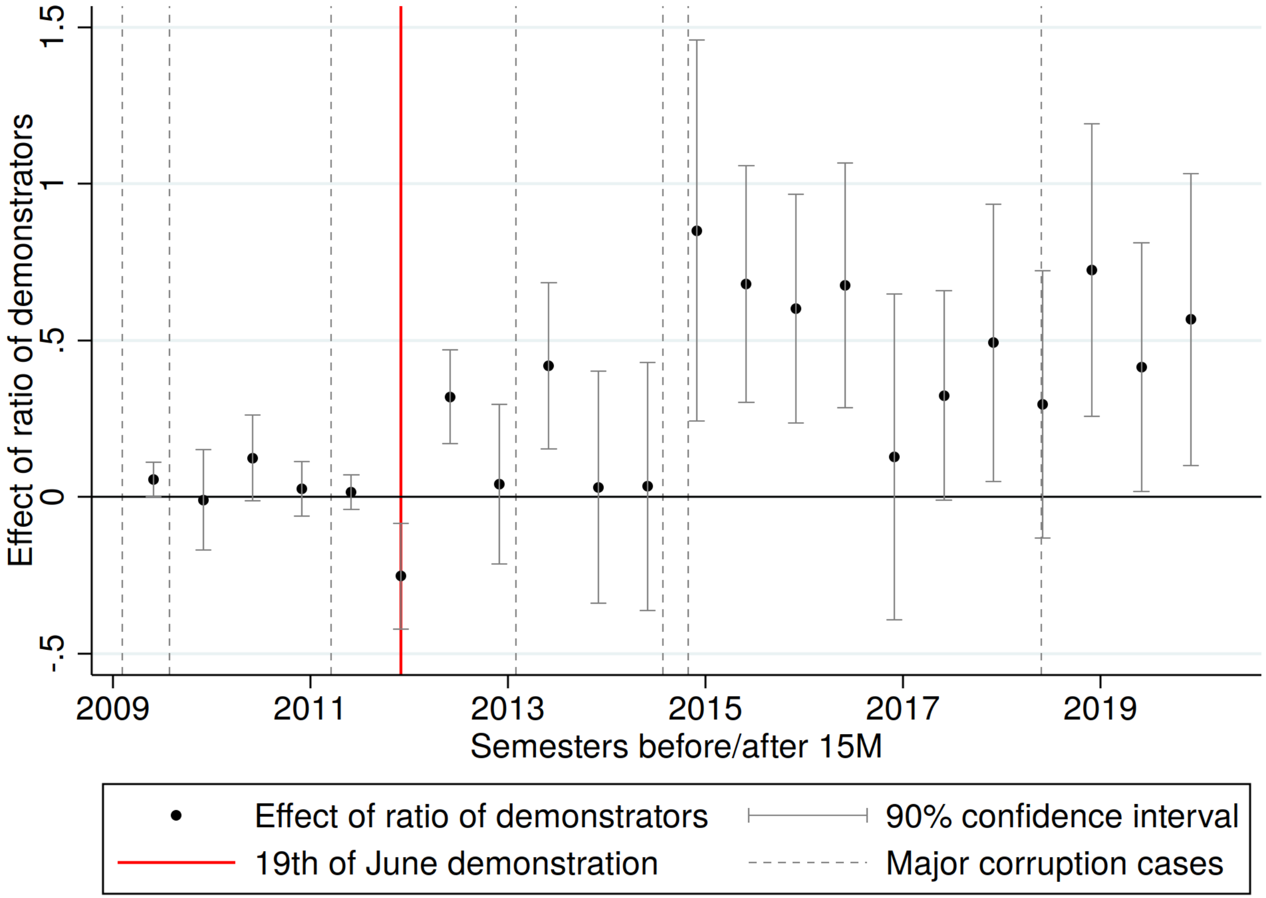
Note: Localisation of all demonstrations of June 19, 2011. Ratio of demonstrators predicted using the first stage regression presented in Table 1. Different colours represent different predicted ratios of demonstrators to city population. Darker colours mean higher predicted proportion of demonstrators. Number of events for each range of predicted participants are in parenthesis.

Figure 4: Evolution of the proportion of people that think that corruption is the top problem that affects Spain



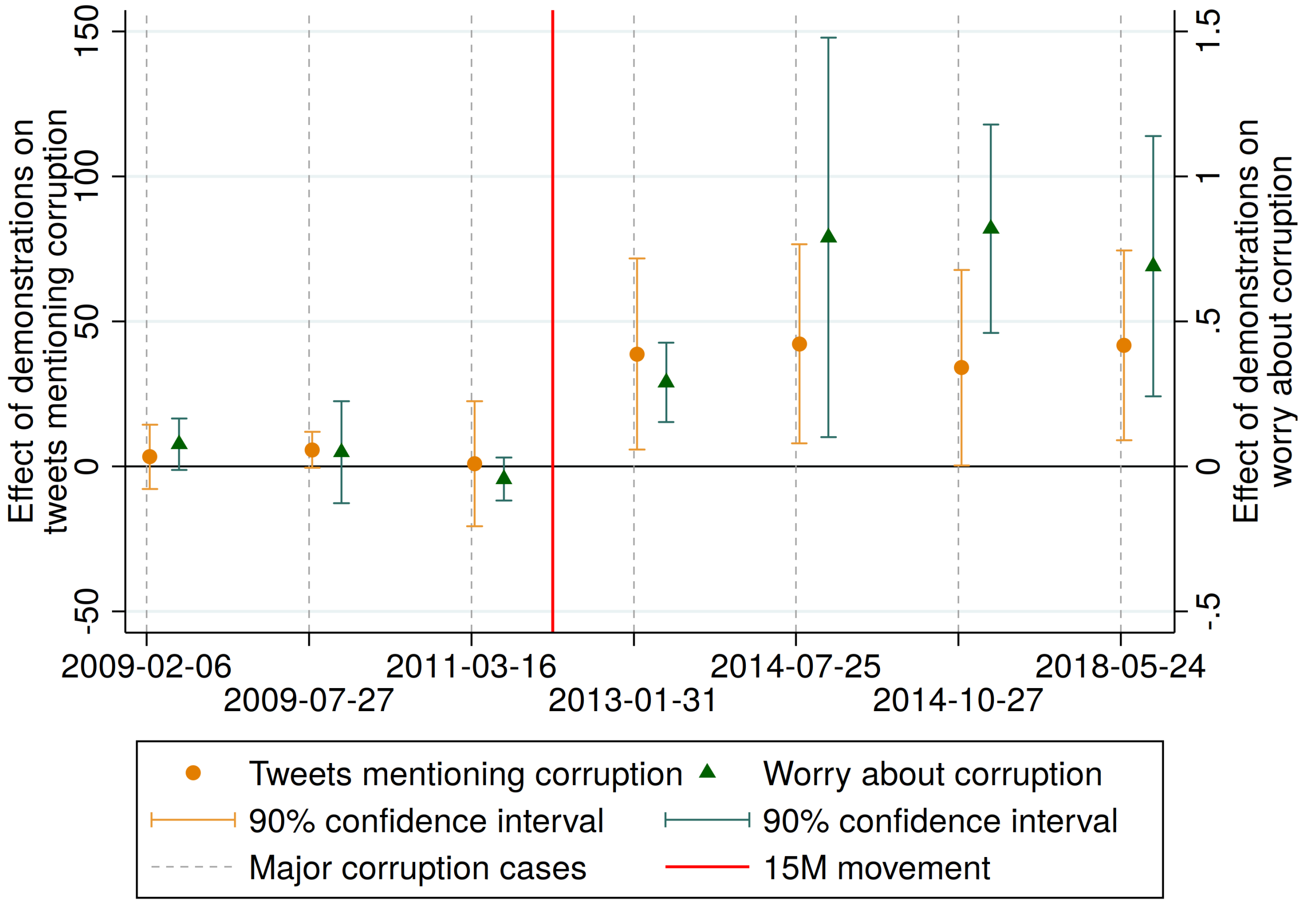
Note: Comparison of the fraction of the population believing that corruption is the top problem affecting Spain in regions with pleasant (hollow circles) and unpleasant (filled triangles) weather during the June 19 protest. The vertical red line represents the June 19 protest, dashed lines represent events of major corruption cases. Data on opinion about corruption concern come from a representative monthly survey from Centro de Investigaciones Sociológicas. Data are pooled by semester, from December to May and June to November.

Figure 5: Estimated effect of the share of protesters in the population (instrumented by unpleasant weather) on concern about corruption over time.



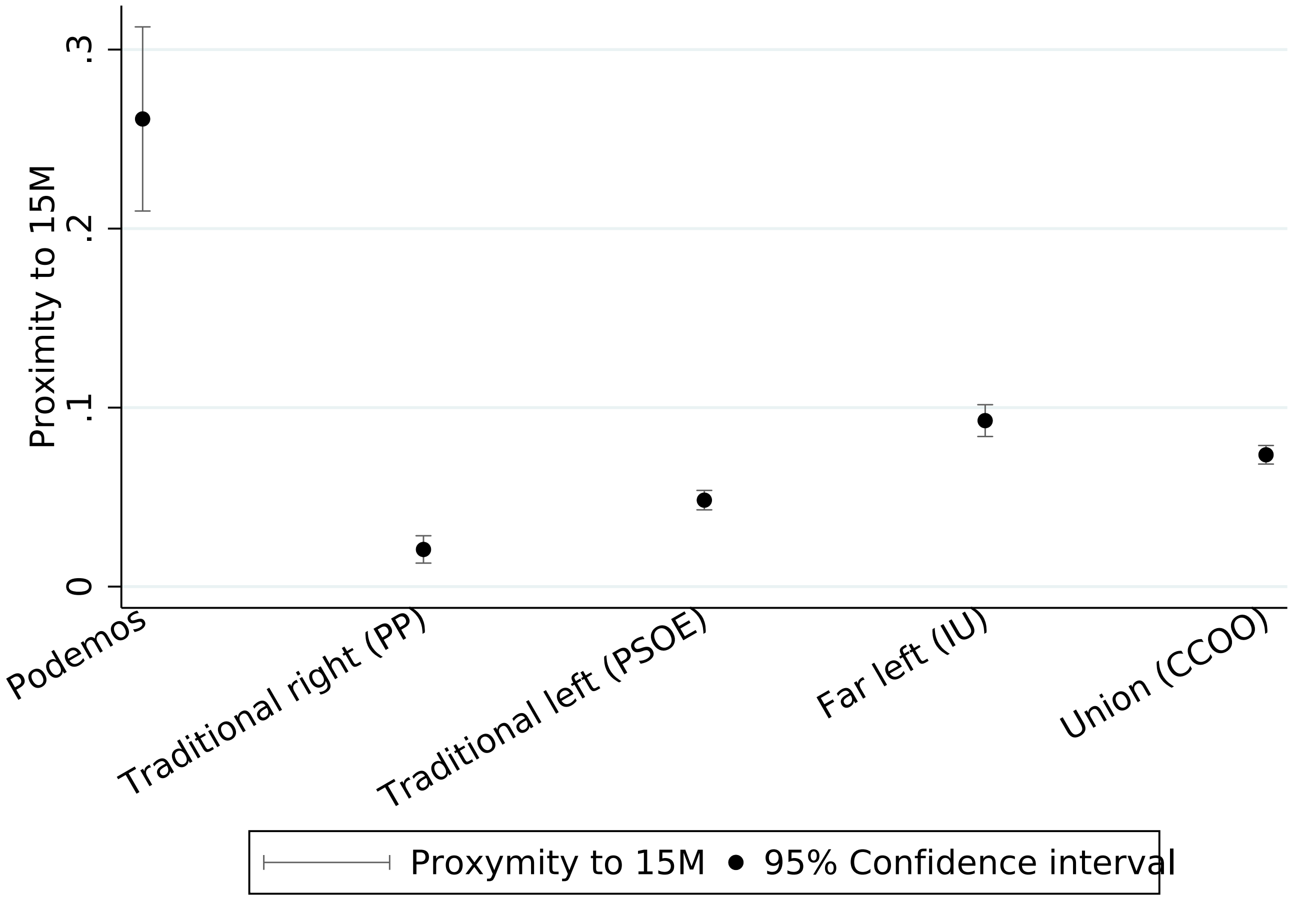
Note: Effect of (instrumented) ratio of demonstrators in the June 19 protest on the probability that an individual declares corruption is the first problem of Spain in the given semester. The vertical red line represents the June 19 protest, dashed lines represent events of major corruption cases. Data about corruption concern come from a representative monthly opinion survey from Centro de Investigaciones Sociológicas. Data are pooled by semester, from December to May and June to November. Regressions include province and individual controls as in Table 3. Confidence intervals shown are 90% confident intervals, standard errors are clustered at the province level.

Figure 6: Estimated effect of the share of protesters in the population (instrumented by unpleasant weather) on tweets and concern about corruption after major corruption cases.



Note: Effect of (instrumented) ratio of demonstrators in the June 19 protest on the number of tweets mentioning corruption in the week following major corruption cases, and on the probability that an individual declares corruption is the first problem of Spain in the semester following major corruption cases. The vertical red line represents the June 19 protest, dashed lines represent events of major corruption cases. Data about corruption concern come from a representative monthly opinion survey from Centro de Investigaciones Sociológicas and are pooled by semester, from December to May and June to November. The underlying regressions are presented in Table 9. Confidence intervals shown are 90% confident interval, standard errors are clustered at the province level.

Figure 7: Proximity of Facebook pages of political organizations to pages related to 15M



Note: Proximity between the network of likes starting from pages related to 15M, and the network of likes of pages of different political parties. Proximity is measured by constructing a set of pages that are liked directly by 15M-related Facebook accounts or liked by pages liked by 15M-related pages, and a similar set from the page of some organization. Proximity represents the proportion of pages in the set of pages liked by 15M that are also in the set of pages liked by the organization. The confidence intervals are computed by considering that we estimate, on a sample of pages liked by pages related to 15M, the proportion of pages that are liked by the other page.

Table 1: First stage regression.

	Ratio demonstrators				
	(1)	(2)	(3)	(4)	(5)
Unpleasant weather	-0.0228***	-0.0215***	-0.0221***	-0.0208***	-0.0214***
	(0.00701)	(0.00687)	(0.00674)	(0.00663)	(0.00654)
Probability of unpleasant		-0.00921	-0.115	0.0113	-0.0842
weather (excl. 2011)		(0.0228)	(0.0992)	(0.0241)	(0.103)
... squared			0.103		0.0923
			(0.0868)		(0.0882)
Probability of unpleasant				-0.0473**	-0.0449*
weather (June 2011)				(0.0232)	(0.0238)
Non-weather city-level controls	Y	Y	Y	Y	Y
F statistic	10.58	9.774	10.71	9.846	10.69
R^2	0.313	0.315	0.328	0.341	0.351
Observations	89	89	89	89	89

Note: Regression of the ratio of demonstrators in the June 19 demonstration over the population of the city on unpleasant weather during this same demonstration. All models includes non-weather city-level controls: population, unemployment rates and the percentage of intention to vote for different parties just before the protest. Additionally, column 2 includes a control for the expected probability of unpleasant weather in June computed from years not including 2011, column 3 adds a squared term for this probability, column 4 instead adds a control for the probability of observing unpleasant weather in June 2011, but excluding three days before and three days after the protest. Finally, column 5 presents the main specification combining all these controls. Robust standard errors in parentheses (*10%, **5%, ***1%).

Table 2: Main table: Electoral effects of protest attendance

Outcomes	Left-right axis				Corruption		
					Anti-	Not anti-	
	Left	Center	Right	Far-right	corruption	corruption	Abstention
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Panel A: 2011 and after						
Ratio demonstrators	1.03**	0.062	-0.57	-0.29**	0.86***	-0.33	-0.53
	(0.47)	(0.12)	(0.52)	(0.14)	(0.32)	(0.60)	(0.61)
Mean of dependent variable	0.306	0.042	0.289	0.034	0.123	0.536	0.341
Observations	623	623	623	623	623	623	623
	Panel B: 2014 and after						
Ratio demonstrators	1.05**	0.11	-0.61	-0.35**	0.97***	-0.42	-0.55
	(0.48)	(0.12)	(0.52)	(0.16)	(0.36)	(0.63)	(0.62)
Mean of dependent variable	0.309	0.043	0.276	0.040	0.132	0.519	0.349
Observations	534	534	534	534	534	534	534
	Panel C: 2019 and after						
Ratio demonstrators	1.19**	0.011*	-0.97	-0.69**	0.42*	-0.24	-0.17
	(0.54)	(0.0068)	(0.65)	(0.31)	(0.25)	(0.63)	(0.61)
Mean of dependent variable	0.326	0.001	0.332	0.076	0.105	0.576	0.319
Observations	267	267	267	267	267	267	267
Election fixed effects	Y	Y	Y	Y	Y	Y	Y
City-level controls	Y	Y	Y	Y	Y	Y	Y

Note: Effect of the ratio of demonstrators in a city, instrumented by unpleasant weather, on aggregated electoral outcomes, pooled over elections. Column 1-4 aggregate all parties along a left-right axis. Respectively, column 1 includes all left-wing parties, column 2 centrist parties, column 3 right-wing parties (including far-right) and column 4 far-right parties. Outcomes of columns 5 and 6 aggregate parties along an anti-corruption / not anti-corruption axis. Column 7 shows results for abstention. Panel A pools together all elections starting from 2011 (2011, 2015, 2016, April 2019 and November 2019 general elections, and 2014 and May 2019 European elections). Panel B pools only the elections starting in 2014, and panel C only the 2019 elections. All models include city-level controls: probability of unpleasant weather in June, variance of unpleasant weather in June, probability of unpleasant weather in June 2011 excluding three days before and three days after the day of the demonstration, population, unemployment rates and the percentage of intention to vote for different parties just before the protest. All models also include election fixed effects. Standard errors (in parentheses) are clustered at the municipality level (*10%, **5%, ***1%).

Table 3: Individual level results: Effect of protest attendance on political opinions

OUTCOMES	Corruption main worry	Political self-classification			
		Left	Right	Far-right	Don't know
	(1)	(2)	(3)	(4)	(5)
		Panel A: 2011 and after			
Ratio demonstrators	0.419***	1.578**	-0.0520	-0.273**	-1.526*
	(0.125)	(0.763)	(0.276)	(0.109)	(0.801)
F first stage	36.22	35.27	35.27	35.27	35.27
Mean of dependent variable	0.121	0.641	0.243	0.0299	0.116
Observations	285,356	258,525	258,525	258,525	258,525
		Panel B: 2014 and after			
Ratio demonstrators	0.501***	1.617**	-0.344	-0.292**	-1.273
	(0.148)	(0.760)	(0.287)	(0.115)	(0.787)
F first stage	35.07	34.11	34.11	34.11	34.11
Mean of dependent variable	0.138	0.644	0.244	0.0309	0.112
Observations	218,314	200,727	200,727	200,727	200,727
		Panel C: 2019 and after			
Ratio demonstrators	0.507**	1.479**	-0.394	-0.345***	-1.084
	(0.220)	(0.659)	(0.304)	(0.129)	(0.728)
F first stage	28.32	27.69	27.69	27.69	27.69
Mean of dependent variable	0.0892	0.646	0.259	0.0369	0.0944
Observations	80,649	75,898	75,898	75,898	75,898
Month fixed effects	Y	Y	Y	Y	Y
City-level controls	Y	Y	Y	Y	Y
Individual-level controls	Y	Y	Y	Y	Y

Note: Effect of the ratio of demonstrators in a city, instrumented by unpleasant weather, on concerns about corruption and on self-placement on the left-right axis (1-10). The sample consists of all individuals interviewed in the monthly "CIS Barometer" between June 2011 and December 2019. Column 1 shows results for responding that corruption is the main problem of Spain. Column 2-4 show the self-placement on the 1-10 left-right axis (left, right, far-right respectively). Column 5 shows the results for the "don't know" answer to the political self-placement question. Panel A pools together all interviews conducted from 2011. Panel B pools interviews from 2014 on, and panel C only interviews conducted during 2019. All models include the usual city-level controls and individual controls including: age, gender, level of studies and unemployment status. Survey-wave fixed effects are also included (a new wave is conducted every month). Standard errors (in parentheses) are clustered at the province level (*10%, **5%, ***1%).

Table 4: Heterogeneous effects of higher protest attendance on electoral results depending on previous political preferences

OUTCOMES	Left-right axis				Corruption		
	Left (1)	Center (2)	Right (3)	Far-right (4)	Anti-corruption (5)	Not anti-corruption (6)	Abstention (7)
	Panel A: Reduced form regression						
Pleasant weather	0.022**	0.0013	-0.012	-0.0063**	0.018***	-0.0070	-0.011
	(0.010)	(0.0026)	(0.011)	(0.0026)	(0.0053)	(0.013)	(0.014)
	Panel B1: Interaction with vote for the left in 2009						
Pleasant weather ×	0.18*	0.021	-0.23	0.021	0.24*	-0.27	0.029
Vote left in 2009	(0.10)	(0.041)	(0.29)	(0.036)	(0.13)	(0.30)	(0.31)
	Panel B2: Interaction with intention to vote for the left in 2011						
Pleasant weather ×	0.12	-0.021	-0.20	-0.010	0.10*	-0.19	0.086
Intention to vote left in 2011	(0.14)	(0.023)	(0.16)	(0.020)	(0.060)	(0.15)	(0.15)
	Panel C1: Interaction with vote for the right in 2009						
Pleasant weather ×	0.28	-0.032	-0.25***	-0.051**	0.062	-0.11	0.044
Vote right in 2009	(0.22)	(0.029)	(0.084)	(0.023)	(0.12)	(0.18)	(0.22)
	Panel C2: Interaction with intention to vote for the right in 2011						
Pleasant weather ×	0.24**	-0.010	-0.044	-0.057***	0.026	0.18	-0.20*
Intention to vote right in 2011	(0.10)	(0.016)	(0.13)	(0.019)	(0.072)	(0.13)	(0.11)
Observations	623	623	623	623	623	623	623
Mean of dependent variable	0.306	0.042	0.289	0.034	0.123	0.536	0.341
Election fixed effects	Y	Y	Y	Y	Y	Y	Y
Probability of unpleasant weather	Y	Y	Y	Y	Y	Y	Y
Population (province)	Y	Y	Y	Y	Y	Y	Y
Unemployment (proxy)	Y	Y	Y	Y	Y	Y	Y
Voting intentions before 15M	Y	Y	Y	Y	Y	Y	Y

Note: Differential effects of pleasant weather interacted with the intention to vote for left or right-wing political options on aggregated electoral outcomes, pooled over elections. Column 1-4 aggregate all parties along a left-right axis. Respectively, column 1 includes all left-wing parties, column 2 centrist parties, column 3 right-wing parties (including far-right) and column 4 far-right parties. Outcomes of columns 5 and 6 aggregate parties along an anti-corruption / not anti-corruption axis. Column 7 shows results for abstention. Panel A shows the reduced form estimate of the effect of pleasant weather. Panels B1 and B2 show the interaction with the vote for the left: panel B1 uses the vote for the left in the 2009 European elections in the municipality. and panel B2 uses the intention to vote for the left from the CIS monthly survey conducted in the year before the 15M movement. Similarly, panel C1 and C2 show the interaction with the vote for the right, with panel C1 using the vote in the 2009 European elections and panel C2 survey data. The sample consists of all cities with demonstrations, and all elections from 2011 to 2019. The models include the usual city-level controls as well as election fixed effects. Standard errors (in parentheses) are clustered at the municipality level (*10%, **5%, ***1%).

Table 5: Heterogeneous effects of higher protest attendance on electoral results depending on age and studies

Outcomes	Corruption main worry	Political self-classification			
		Left	Right	Far-right	Don't know
	(1)	(2)	(3)	(4)	(5)
	Panel A: Reduced form regression				
Pleasant weather	0.0149***	0.0560**	-0.00185	-0.00971***	-0.0542*
	(0.00464)	(0.0254)	(0.00990)	(0.00361)	(0.0271)
	Panel B: Interaction with young				
Pleasant weather × Young	-0.00946*	0.0422**	-0.0265***	0.00314	-0.0158
	(0.00494)	(0.0178)	(0.00717)	(0.00255)	(0.0161)
	Panel C: Interaction with higher studies				
Pleasant weather × Higher studies	-0.00267	0.0464**	-0.0611***	-0.00432	0.0147
	(0.00228)	(0.0189)	(0.0172)	(0.00299)	(0.0233)
Mean of dependent variable	0.121	0.641	0.243	0.0299	0.116
Observations	285,356	258,525	258,525	258,525	258,525
Month fixed effects	Y	Y	Y	Y	Y
City-level controls	Y	Y	Y	Y	Y
Individual-level controls	Y	Y	Y	Y	Y

Note: Effect of interaction of pleasant weather during the June 19, 2011 demonstration and individual characteristics on concerns about corruption and on self-placement on the left-right axis (1-10). The sample consists of all individuals interviewed in the monthly "CIS Barometer" between June 2011 and December 2019. Column 1 shows results for responding that corruption is the main problem of Spain. Column 2-4 show the self-placement on the 1-10 left-right axis (left, right, far-right respectively). Column 5 shows the results for the "don't know" answer to the political self-placement question. Panel A shows the reduced form estimate of the effect of pleasant weather, Panel B the interaction with a dummy indicating whether the individual was 20 years old or less in 2011, and Panel C the interaction with a dummy indicating whether the individual completed a higher education curriculum. All models include the usual city-level controls and individual controls as well as month fixed effects. Standard errors (in parentheses) are clustered at the province level (*10%, **5%, ***1%).

Table 6: Twitter activity before and after the movement

OUTCOMES	First principal component (1)	Log(Tweets) (2)	Log(Users) (3)	Log(New users) (4)
	Panel A: Placebo (April 2011)			
Ratio demonstrators	-3.29	-0.99	7.69	-12.82
	(21.70)	(19.62)	(19.31)	(17.48)
Mean of dependent variable	0.000	5.070	4.329	1.886
	Panel B: 19th of June to 4th of July 2011			
Ratio demonstrators	43.51**	42.39**	40.58**	21.43**
	(19.37)	(19.04)	(18.05)	(10.42)
Mean of dependent variable	0.000	5.141	4.473	1.232
City-level controls	Y	Y	Y	Y
Observations	89	89	89	89

Note: Effect of the ratio of demonstrators in a city, instrumented by unpleasant weather, on measures of Twitter activity in the city. The measures are derived from a random sample of Spanish-language tweets in a 14 days period collected by searching for the 100 most frequent words in Spanish. Panel A uses a placebo sample in April 2011. Panel B uses a sample of tweets posted between the 19th of June and the 4th of July. Column 2 shows as outcome the logarithm of the number of tweets located to the given city, column 3 the number of distinct users from the city observed in the sample, and column 4 the number of new users (created after the beginning of the sample collection). Column 1 shows the effect on the first principal component derived from these three measures. Its composition is detailed in Table D3. All models include the usual city-level controls. Robust standard errors in parentheses (*10%, **5%, ***1%).

Table 7: Tweets after protests mentioning 15M-associated words

OUTCOMES	Tweets with hashtags associated with 15M posted between June 19 and July 19, 2011		
	Log(Tweets)	Log(Users)	Log(Tweets positive)
	(1)	(2)	(3)
Ratio demonstrators	47.36**	42.90**	40.37*
	(23.81)	(18.84)	(21.44)
Mean of dependent variable	4.925	3.577	3.085
Corruption scandal fixed effects	Y	Y	Y
City-level controls	Y	Y	Y
Observations	89	89	89

Note: Effect of the ratio of demonstrators in a city, instrumented by unpleasant weather, on measures of 15M-related Twitter activity in the city. The measures are derived from collecting all tweets containing hashtags and words related to 15M between June 19 and July 19, 2011. The list of keywords is presented in Table D2. Column 1 shows as outcome the logarithm of the number of tweets located in the given city, column 2 the number of distinct users from the city observed in the sample, and column 3 the tweets expressing a positive sentiment (see §D.3). All models include the usual city-level controls. Robust standard errors in parentheses (*10%, **5%, ***1%).

Table 8: Pre-election Twitter activity related to 15M and Podemos

OUTCOMES	Tweets mentioning 15M		Tweets mentioning Podemos	
	Log(Tweets) (1)	Log(Users) (2)	Log(Tweets) (3)	Log(Users) (4)
	Panel A: before 2015 election			
Ratio demonstrators	53.70**	43.18**	69.64**	67.35***
	(21.53)	(17.98)	(28.66)	(26.05)
Mean of dependent variable	2.642	2.249	6.284	5.100
	Panel B: before April 2019 election			
Ratio demonstrators	52.73***	45.74**	74.04**	68.15**
	(20.09)	(17.83)	(29.27)	(27.26)
Mean of dependent variable	1.966	1.798	6.541	5.378
City-level controls	Y	Y	Y	Y
Observations	89	89	89	89

Note: Effect of the ratio of demonstrators in a city, instrumented by unpleasant weather, on measures of 15M and Podemos-related Twitter activity in the city before elections. The measures are derived from collecting all tweets mentioning 15M or Podemos in the given period. Panel A shows tweets posted during the month before the December 20, 2015 election, panel B tweets posted during the month before the April 28, 2019 election. Column 1 shows as outcome the logarithm of the number of tweets mentioning 15M located in the given city, column 2 the logarithm of the number of distinct users mentioning 15M from the city, and column 3 and 4 the logarithm of the number of tweets and users mentioning Podemos. Tweets were collected on August 20, 2021. All models include the usual city-level controls. Robust standard errors in parentheses (*10%, **5%, ***1%).

Table 9: Tweets and concerns about corruption after corruption scandals

	Corruption case			
	2013-01-31 (Barcenas) (1)	2014-07-25 (Pujol) (2)	2014-10-27 (Punica) (3)	2018-05-24 (Barcenas) (4)
	Panel A: Effect on Log(Tweets)			
Ratio demonstrators	38.74*	42.25**	34.04*	41.67**
	(20.01)	(20.88)	(20.50)	(19.90)
Mean of dependent variable	5.210	3.947	5.723	5.983
Observations	89	89	89	89
	Panel B: Effect on Log(Users)			
Ratio demonstrators	37.99**	37.89**	33.78*	38.74**
	(18.42)	(17.92)	(17.80)	(18.26)
Mean of dependent variable	4.564	3.371	4.579	4.674
Observations	89	89	89	89
	Panel C: Effect on corruption concern			
Ratio demonstrators	0.289***	0.790*	0.820***	0.690**
2SLS standard error	(0.0835)	(0.419)	(0.219)	(0.273)
Individual-level controls	Y	Y	Y	Y
Month fixed effects	Y	Y	Y	Y
Mean of dependent variable	0.0335	0.223	0.226	0.142
F first stage	38.19	38.51	38.69	39.64
Observations	14,837	12,280	14,785	13,794
Corruption scandal fixed effects	Y	Y	Y	Y
City-level controls	Y	Y	Y	Y

Note: Effect of the ratio of demonstrators in a city, instrumented by unpleasant weather, on measures of corruption-related Twitter activity and corruption concerns. Each column corresponds to an important date in a major corruption scandal. Panel A and B show the effects on Twitter activity. The measures are derived from collecting all tweets mentioning corruption in the week following the revelation of new information. Panel A uses as outcome the logarithm of the number of tweets mentioning corruption, while Panel B presents the logarithm of the number of distinct users posting during the period. Panel C uses individual-level survey data from the CIS barometer from the six months following the indicated date, and the outcome is a dummy variable equal to one if the respondent's main political concern is corruption. All models include the usual city-level controls. Additionally, panel C includes the usual individual-level controls and month fixed effects. Standard errors (in parentheses) are robust to heteroscedasticity in Panels A and B, and clustered at the province level for panel C (*10%, **5%, ***1%).

Table 10: Differential effect of having a social media account on electoral choices

Outcomes	Left-right axis				Corruption		
	Left (1)	Center (2)	Right (3)	Far-right (4)	Anti-corruption (5)	Not anti-corruption (6)	Abstention (7)
	Panel A: 2011 and after						
Pleasant × Social media account	0.0687***	-0.0277**	-0.0261*	0.00214	0.0244***	-0.0104	-0.0140
	(0.0142)	(0.0116)	(0.0142)	(0.00585)	(0.00770)	(0.0127)	(0.0105)
Observations	36,923	36,923	36,923	31,447	36,923	36,923	36,923
	Panel B: 2014 and after						
Pleasant × Social media account	0.0714***	-0.0296**	-0.0237*	0.00214	0.0263***	-0.0110	-0.0153
	(0.0132)	(0.0122)	(0.0140)	(0.00585)	(0.00842)	(0.0138)	(0.0107)
Observations	31,447	31,447	31,447	31,447	31,447	31,447	31,447
	Panel C: 2019 and after						
Pleasant × Social media account	0.0654***	-0.0276*	-0.0267	-0.00246	0.0144	0.00937	-0.0238
	(0.0161)	(0.0142)	(0.0173)	(0.0117)	(0.0113)	(0.0218)	(0.0179)
Observations	16,979	16,979	16,979	16,979	16,979	16,979	16,979
Election fixed effects	Y	Y	Y	Y	Y	Y	Y
City-level controls	Y	Y	Y	Y	Y	Y	Y
Individual controls	Y	Y	Y	Y	Y	Y	Y

Note: Effect of the interaction between pleasant weather during the June 19, 2011 demonstration in the capital of a province, instrumented by unpleasant weather, and an individual in the province having a social media account, on aggregated electoral outcomes, pooled over elections. The sample comes from the CIS post-electoral surveys for all elections from 2011 to 2019. Column 1-4 aggregate all parties along a left-right axis. Respectively, column 1 includes all left-wing parties, column 2 centrist parties, column 3 right-wing parties (including far-right) and column 4 far-right parties. Outcomes of columns 5 and 6 aggregate parties along an anti-corruption / not anti-corruption axis. Column 7 shows results for abstention. Panel A pools together all elections starting from 2011 (2011, 2015, 2016, April 2019 and November 2019 general elections, and 2014 and May 2019 European elections). Panel B pools only the elections starting in 2014, and panel C only the 2019 elections. All models include the usual city-level, individual-level controls, and election fixed effects. Standard errors (in parentheses) are clustered at the province level. (*10%, **5%, ***1%)

Online Appendix

Contents

A Robustness checks

A.1 Robustness of the first stage

While weather is plausibly exogenous, it does not fully obey the exclusion restriction: weather is known to impact a wide range of economic and social phenomena, including income and conflict. While this study only relies on the weather on the day of the demonstration, this does not entirely remove all possible causal links. Even if the 19th of June demonstrations can be considered have been the most important event on that day in Spain (as can be confirmed, for instance, by observing the front page of newspapers the next day), it is not the only event happening on that day that might have had political relevance. Moreover, weather is temporally correlated: weather differences between cities on the 19th of June might be correlated with differences during a wider period, and the differences during a wider period are more likely to have impacts through other channels. I deal with this issue in multiple steps.

Adding weather in surrounding months as a control The main specification already includes a control for the probability of unpleasant weather occurring in June 2011, excluding the days surrounding the demonstration. This is already an indication that the effect is driven by the weather around the time of the demonstration and not the longer term weather patterns that year. To gain more assurance of this, Table A1 reports the result of the main regression also controlling for the number of pleasant days in May and July. The estimates show very little variation, and stay significant: this shows that the results are driven by the local weather shocks around the 19th of June and not the wider weather patterns in the surrounding months.[38] Thus, violations of the exclusion restriction are limited to direct effects of the weather in the week around the 19th of June. In particular, this implies that the variation we explain is unlikely to be explained by the weather on election days 5 months to 8 years later. This also greatly limit the magnitude of potential violations: for example, while weather can explain internal migration (Rappaport, 2009), the weather in a particular week of 2011 will only play a proportionally tiny role in a longer-term phenomenon.

Sensitivity analysis In a next step, I follow Mellon (2021) and conduct a sensitivity analysis (Cinelli and Hazlett, 2020b) to evaluate how large an exclusion restriction violation would have to be to render the results insignificant. Table A3 presents sensitivity measures for our results. For each panel, we report the partial R^2 of the instrument in the reduced for regression, and the robustness value $RV_{q,\alpha}$ of Cinelli and Hazlett (2020b) for a significance level of 0.1: the results indicate that to overturn the result, an exclusion restriction violating variable Z would have to have a partial R^2 with weather of at least

[38]The corresponding first stage is presented in Table A2.

this value, and a partial R^2 with the outcome above this value. Outcomes more strongly correlated with the weather would need to be less strongly correlated with the electoral outcomes and conversely.

To get a less abstract evaluation, I examine some plausible channels for violations of the exclusion restriction using partial correlations that can be found in the literature. I reuse the values from (Mellon, 2021) directly, and study the possible effects of violent protest, property crime, pollution and income: the level of violence during protests is related to temperature (Wasow, 2020) and violent protests might have more exposure or push citizens away from the movement; property crime might influence elections by pushing voters to the right; pollution has impacts on the wide variety of outcome including health, mortality, etc ; income may influence voters choices (Mellon, 2021). For each of these variables and each significant outcome, I show in Table A3 the minimum partial R^2 value of this variable with the outcome that would be necessary to make the results non-significant. The results for left-wing votes are very robust: violent protest and property crime around the 19th of June would have to explain more than 40% of the variation in electoral outcomes up to 8 years later to overturn the results. The other results require R^2 that are up to an order of magnitude smaller. We can use the fact that we are only depending on weather during a very small interval to argue that many of these causes actually cannot overturn our results. Consider for example the effect of pollution on the 2011 result for the far-right vote. To overturn the result, pollution caused by the weather around the 19th of June should explain 20.1% of the remaining unexplained part of the electoral results. But there are 22 weeks between the 19th of June protest and the 2011 election, and it is likely that pollution in each of these weeks would have a similar effect. Then, the variation of pollution over the whole period before the election would have to essentially explain far more variation than is actually present in the electoral result. The other results are even stronger, since the time period involved is larger, and the effect of pollution on one day become even more diluted.

Generally speaking, following the same reasoning, violations of the exclusion restriction that could plausibly overturn the results would need to be particularly linked to the precise date of the demonstration, or the week around it. This is the case for confounders that are related to the 15M movement or the demonstration itself. For instance, more pleasant weather on the day of the march may be associated to a higher presence of journalists to cover the march, leading to a more favorable media coverage. Even if, in all likelihood, journalists may have decided whether to be present the previous day, working in better conditions might improve the coverage. Similarly, weather has an effect on the mood of participants: it is likely that participants feeling happier during the demonstration would return with more positive associations with the movement, which might influence their future participation. Finally, the weather between the 16th and the 22nd of June is likely to be related to the success of the camps created by the movement in squares around Spain, although this would only be a short time in a camp's existence. If these causal channels play a significant role, they would not fully invalidate the conclusions of this study: the results still show that the success of a protest movement can lead to long-term changes, although the precise measure of success is not necessarily the attendance at the protests.

Removing some weather controls In Table A4, I examine whether the results are sensitive to changing the set of weather controls used in the regression, by looking at the main regression using all elections after 2011. The first stage regressions are the ones presented on Table 1. The F statistic is sometimes slightly under 10, and the instrument can then be considered weak. I thus also present AR test p-values robust to weak instruments. Panel E presents the main specification for reference. In Panel A, I remove all weather controls. In Panel B, I use only the expected weather excluding 2011. In Panel C, I add the squared term. In Panel D, I show results using both the expected probability of unpleasant weather excluding 2011 and the probability in 2011. The results stay consistent between all panels, with similar coefficients and errors. The only exception is for the vote for the left in Panel A: when removing all controls, the coefficient is lower . This indicates that the probability of unpleasant weather does capture some important pre-existing geographic characteristics that influence the vote for the left (on the other hand, the results of the first stage are not significantly different).

Weather does not predict unemployment Unemployment rate is one of the variables that is likely to be related to protest attendance (because the opportunity cost of attending is lower) and to electoral results (because unemployed rate have an impact in electoral results (Artés, 2014; Garz and Martin, 2021)). It a variable particularly relevant for the specific context we study because in 2011, Spain had an unemployment rate higher than 25% and around 40% for people below 25 years old. To test whether unemployment is linked to weather shocks on the day of the march (the treatment variable), I run the first stage regression on unemployment rate instead of the number of demonstrators. The dependent variable is unemployment and the independent variables is unpleasant weather. Consequently, unemployment controls are removed from the equation. Results are in Table A5. As expected we do not observe a significant effect.

Placebo: weather shocks on other days I check that there is no spurious correlation left between weather shocks and the number of demonstrators that could be driving the results. First, I run the first stage using weather conditions during 240 days in the month of June from years 2010 to 2018 (excluding 2011). Results are reported in Figure A1 and show that, on average, weather shocks are not significant in predicting the number of demonstrators on June 19. Relative to the distribution of the placebo coefficients of unpleasant weather in the first stage regression, the value of the coefficient of unpleasant weather during the day of the marches is lower than 97.5% of all the placebo dates, and that the F statistic is higher than 95% of placebo dates.

A.2 Robustness of the main results

Placebo for outcomes before the demonstration I replicate the study using as outcome elections before the demonstration, and present the results in Table A6. Panel A uses as outcome the 2000 congressional elections, panel B the 2004 congressional elections and panel C the 2008 congressional elections. I also consider in Panel B and C the 2004 and 2008 congressional elections only. We expect to observe no effect on outcomes measured before the demonstration. We do observe one significant result: there is a negative effect on the far-right in 2011. However, we are testing 21 different hypothesis here, and it is not surprising to see one result significant at the $p < 0.1$ level.

Similarly, I replicate in Table A7 the study from Table 3, using as outcome responses to the survey from May 2010 to April 2011. Similarly, we observe no statistically significant effect of the number of demonstrators on these outcomes, except for the "don't know" classification where we observe a negative effect.

Excluding one city at the time To show that the results are not driven by a single outlier, I estimate all the significant results excluding one city at a time. Figure A2a shows the results for the first stage. The top panel shows the coefficient with the 90% confidence interval while the bottom panel shows the F statistic. We see that the coefficient remains significant when excluding any city even if Madrid drives an important part of the first stage results. Figures A2 and A3 show the coefficients for the second stage for all significant outcomes (with the 90% 2SLS interval) and the AR p-value (robust to weak instruments). Overall, we see that in general the results remain significant to the exclusion of any singular city, with the exception of anti-corruption vote in 2019 that becomes insignificant (at 10% level) when excluding any city of a group of around a half of the cities. In 2011 and 2014, the exclusion of Madrid still increases the confidence intervals and p-values (probably due to the fact that the instrument is weaker) past the usual thresholds of significance.

Changing the definition of the main independent variable Instead of using the ratio of the number of demonstrators over the population of the municipality, I use the logarithm of the number of demonstrators. The first stage is shown in Table A8. The results for the left and far-right still hold, although with reduced significance (Table A9). We cannot compare their magnitudes as they use a different measure of protest. The results for the anti-corruption vote are stay similarly significant for Panel A and B, but become barely insignificant for the 2019 elections (Panel C, $p = 0.105$).

Adding additional control variables Following Lu and White (2014), I added control variables as a robustness check that can increase precision of the estimates.

Looking at the way different regions of Spain behave electorally, it is clear that Catalonia and the Basque Country stand out from the rest. To account for that, I added two dummy variables, one for Catalonia and one for Basque country. Results are also presented in Panel A of Table A10, and first stage results in column 1 of Table A10. All results remain significant.The magnitude of the results for the far-right and for anti-corruption parties are slightly lower, while the magnitude for the left is slightly higher.

I include province per-capita GDP as a measure of income of the region. Richer regions may vote differently than poorer regions. The results, presented in Panel B of Table A10, are very similar: all significant results stay significant, although the magnitude of the effects is generally lower. The first stage is presented in column 2 of Table A11.

Robustness to weak instruments The F statistic for the first stage (F = 10.69), although above the threshold of 10, indicates that the instrument is not very strong. In Table A12, I additionally present AR p-values robust to weak instruments (Anderson and Rubin, 1949), following (Mikusheva and Poi, 2006). The results are robust to this test.

Accounting for spatial correlation Weather is likely to be spatially correlated and this can affect the size of the standard errors. To account for that I compute standard errors robust to spacial correlation (Conley, 1999, following König et al., 2017). The results are presented for a selection of thresholds (50, 100, 150, and 200km) on Table A13.

The effects stay significant for all distances: the standard errors are generally lower than the standard errors in Table 2, indicating that there may be some negative spatial correlation between nearby cities. For a distance of 200 km for anti-corruption vote in 2019, the result is not significant anymore. Additionally, accounting for spatial correlation shows significant negative results on the right-wing vote for all three time periods.

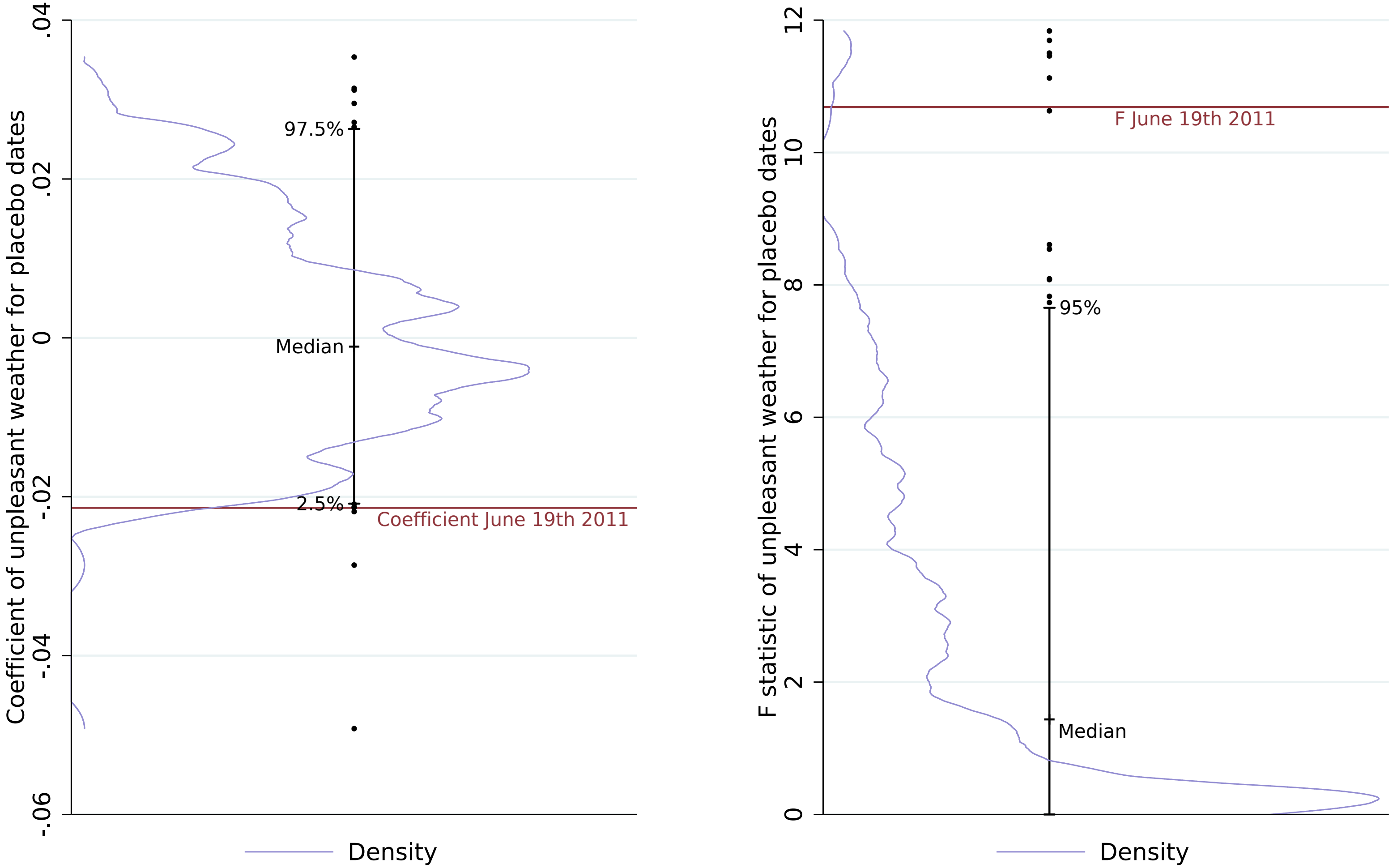
Clustering by province In the general elections, representatives are elected at the province level. Thus, there may be intra-province correlation in electoral results. To control for this, I cluster the standard errors at the province level. The first stage is shown in Table A14, and the second stage regressions are reported in Table A15. The errors are very close to the unclustered errors in the main specification, and all results keep the same significance level. The instrument is weaker when clustering by province: the F statistic of the first stage is 9.13, lower than the usual threshold of 10 (Stock and Yogo, 2005): I also present AR p-values for the estimates. Again, all results stay significant for this test.

Multiple hypothesis testing In Table A16, I use the Romano and Wolf (2005) method to compute p-values adjusted for multiple hypothesis testing from the 2SLS estimate. All results stay significant.

Figure A1: Robustness check. Distribution of coefficients and F-statistics, corresponding to different regressions of unpleasant weather in random days of June in years around 2011 on the logarithm of the number of demonstrators on June 19th, 2011.

(a) Coefficient

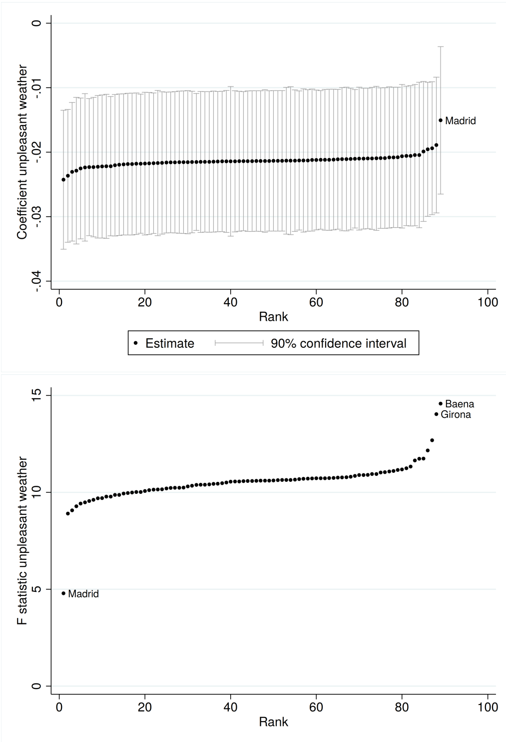
(b) F statistic



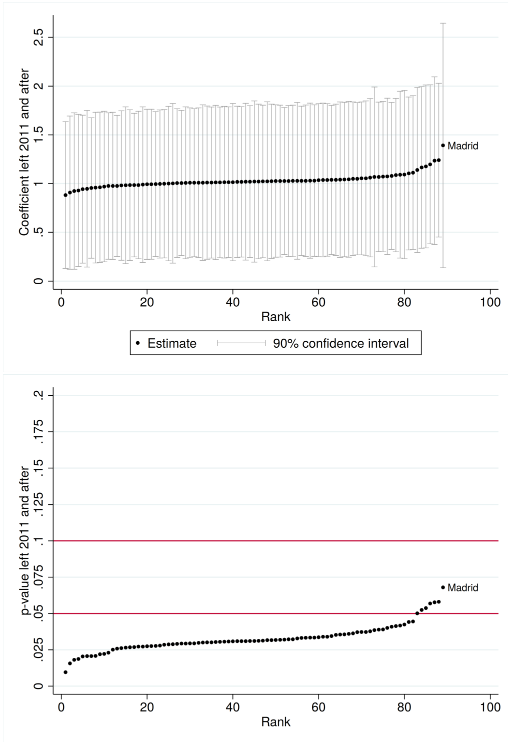
Note: Density plot of the coefficients and F statistic of placebo first stages, using as predictor unpleasant weather on all possible days in June between 2010 and 2018, excluding 2011, controlled by the expected unpleasant weather in June (240 placebo regressions in total). The red line corresponds to the coefficient and F statistic of the regression of unpleasant weather during the 19th June 2011 demonstration on the ratio of the number of demonstrators to the population on this same date (which is the coefficient and the F statistic of the first stage of this study). The median, 2.5-th and 97.5th percentile are displayed for the coefficient, and the 95th percentile for the F statistic. Points outside this interval are displayed as outliers.

Figure A2: Removing one observation at a time

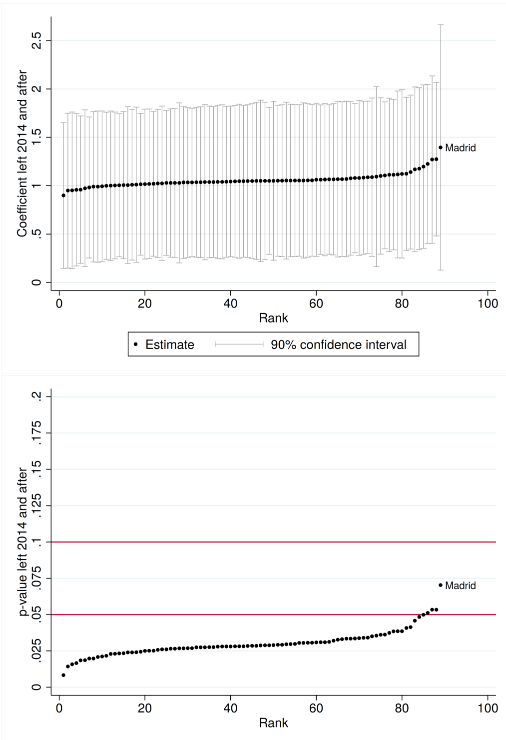
(a) First stage



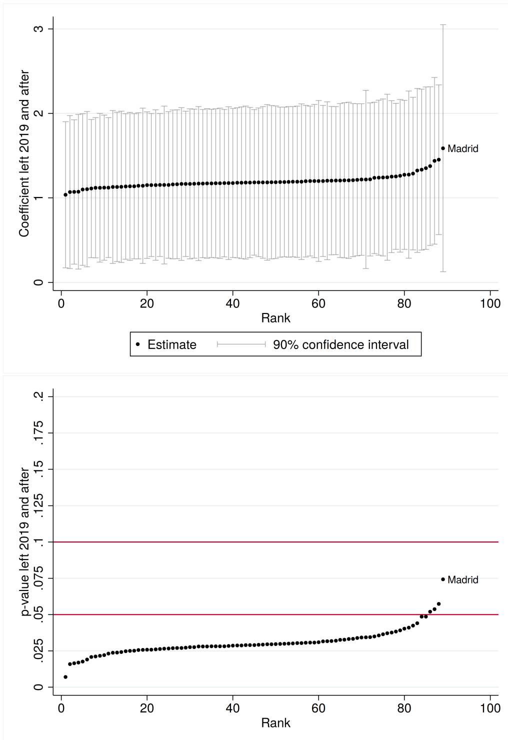
(b) Left-wing parties, 2011 and after



(c) Left-wing parties, 2014 and after



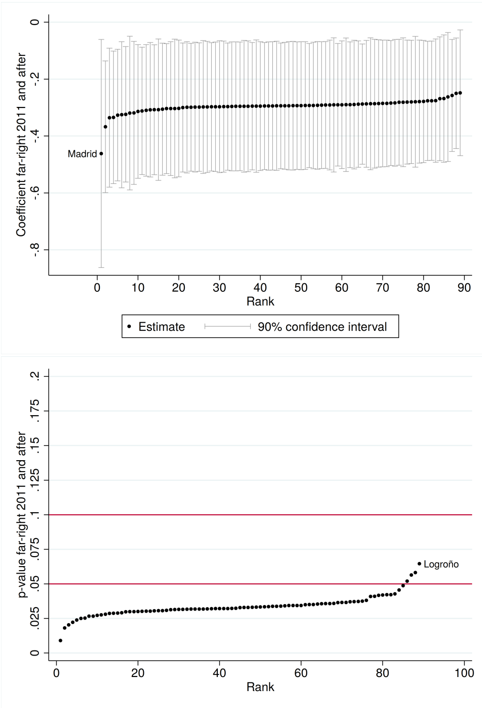
(d) Left-wing parties, 2019



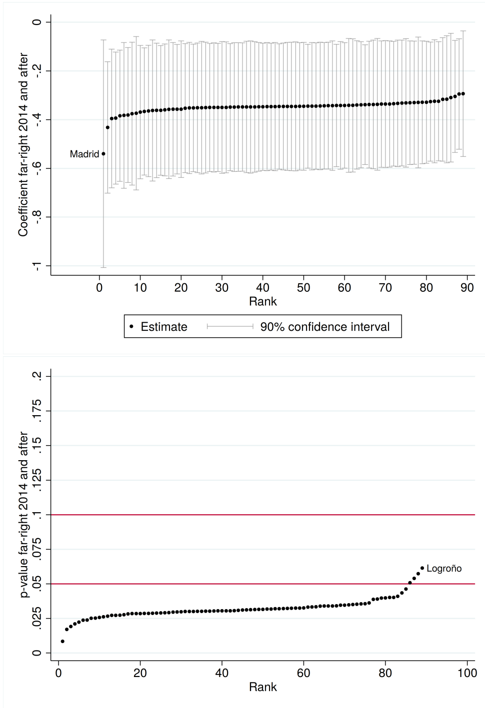
Note: Graphs representing the results of the first and second stage for different outcomes when removing each observation in turn. Each point corresponds to a regression with one specific city removed. Subfigure (a) presents the results of the first stage: the top graph represents the coefficient of unpleasant weather in the first stage along with its 90% confidence interval, the bottom graph the associated F statistic. Other subfigures correspond to different electoral outcomes. The top graph represents the estimate of the effect of the ratio of demonstrators and its 90% confidence interval, the bottom graph represents the p-value. Red horizontal lines represent the $p < 0.1$ and $p < 0.5$ confidence thresholds.

Figure A3: Removing one observation at a time (2)

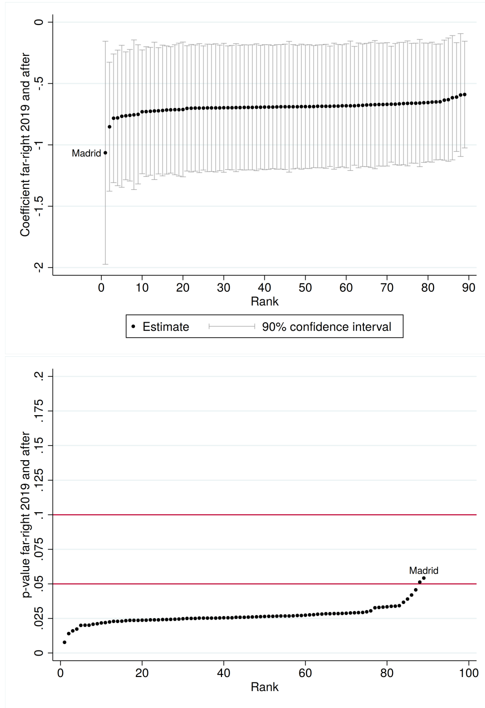
(a) Far-right parties, 2011 and after



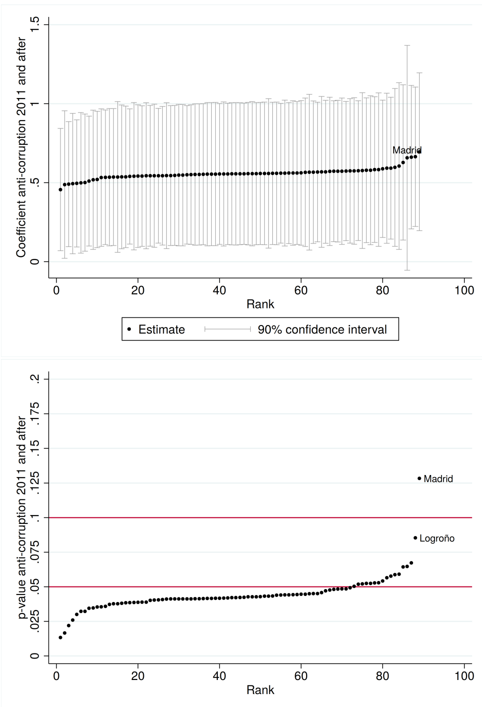
(b) Far-right parties, 2014 and after



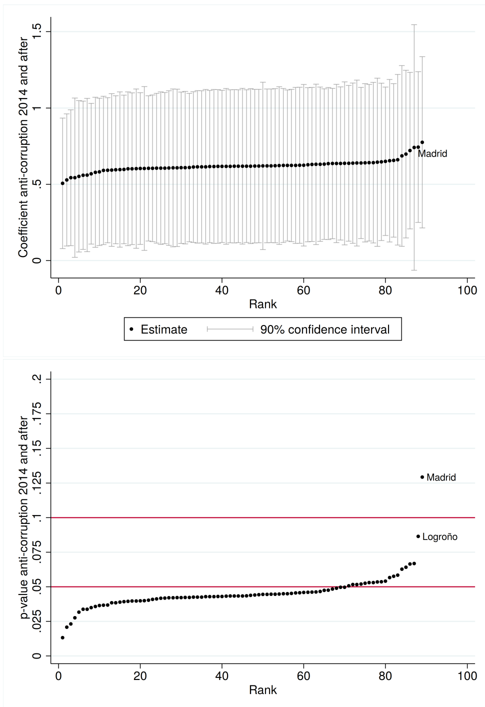
(c) Far-right parties, 2019 and after



(d) Anti-corruption vote, 2011 and after



(e) Anti-corruption vote, 2014 and after



(f) Anti-corruption vote, 2019 and after

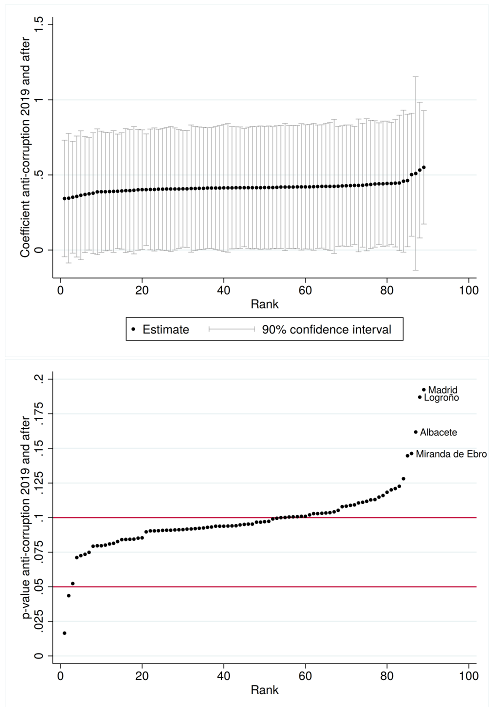


Table A1: Robustness check: controlling by weather in May and July

Outcomes	Left-right axis				Corruption		
	Left (1)	Center (2)	Right (3)	Far-right (4)	Anti-corruption (5)	Not anti-corruption (6)	Abstention (7)
	Panel A: 2011 and after						
Ratio demonstrators	1.08**	0.072	-0.41	-0.29**	0.85***	-0.088	-0.76
	(0.47)	(0.12)	(0.48)	(0.14)	(0.31)	(0.56)	(0.57)
Mean of dependent variable	0.306	0.042	0.289	0.034	0.123	0.536	0.341
Observations	623	623	623	623	623	623	623
	Panel B: 2014 and after						
Ratio demonstrators	1.11**	0.11	-0.44	-0.34**	0.96***	-0.18	-0.78
	(0.48)	(0.12)	(0.48)	(0.16)	(0.35)	(0.58)	(0.58)
Mean of dependent variable	0.309	0.043	0.276	0.040	0.132	0.519	0.349
Observations	534	534	534	534	534	534	534
	Panel C: 2019 and after						
Ratio demonstrators	1.28**	0.012*	-0.80	-0.68**	0.40*	0.023	-0.42
	(0.54)	(0.0069)	(0.61)	(0.31)	(0.24)	(0.57)	(0.56)
Mean of dependent variable	0.326	0.001	0.332	0.076	0.105	0.576	0.319
Observations	267	267	267	267	267	267	267
Weather in June (3 days away)	Y	Y	Y	Y	Y	Y	Y
Weather in May and July	Y	Y	Y	Y	Y	Y	Y
Election fixed effects	Y	Y	Y	Y	Y	Y	Y
City-level controls	Y	Y	Y	Y	Y	Y	Y

Note: Effect of the ratio of demonstrators in a city, instrumented by unpleasant weather, on aggregated electoral outcomes, pooled over elections. Column 1-4 aggregate all parties along a left-right axis. Respectively, column 1 includes all left-wing parties, column 2 centrist parties, column 3 right-wing parties (including far-right) and column 4 far-right parties. Outcomes of columns 5 and 6 aggregate parties along an anti-corruption / not anti-corruption axis. Column 7 shows results for abstention. Panel A pools together all elections starting from 2011 (2011, 2015, 2016, April 2019 and November 2019 general elections, and 2014 and May 2019 European elections). Panel B pools only the elections starting in 2014, and panel C only the 2019 elections. All models include the usual city-level controls as well as include election fixed effects. Additionally, all regressions include the probability of a day in May or July 2011 having pleasant weather. Standard errors (in parentheses) are clustered at the municipality level (*10%, **5%, ***1%).

Table A2: First stage regression, controlling by weather in May and July.

	Ratio demonstrators (1)
Unpleasant weather	-0.0220***
	(0.00666)
Probability of unpleasant weather	0.00895
	(0.0256)
City-level controls	Y
Mean of dependent variable	0.033
F statistic for unpleasant weather	10.89
R^2	0.35
Observations	89

Note: Regression of the ratio of demonstrators in the June 19 demonstration over the population of the city on unpleasant weather during this same demonstration. The model includes the usual city-level controls, as well as a control for the probability of a day in May or July 2011 having pleasant weather. Robust standard errors in parentheses (*10%, **5%, ***1%).

Table A3: Sensitivity analysis of the instrument

Outcomes	Left (1)	Far-right (2)	Anti-corruption (3)
	Panel A: 2011 and after		
Partial R^2	0.042	0.015	0.020
Robustness value	0.132	0.0537	0.0741
... to violent protest	0.565	0.658	0.135
... to property crime	0.584	0.0674	0.139
... to pollution	>1	0.345	>1
... to income	>1	>1	>1
	Panel B: 2014 and after		
Partial R^2	0.045	0.019	0.023
Robustness value	0.134	0.0629	0.0774
... to violent protest	0.61	0.939	0.151
... to property crime	0.635	0.962	0.155
... to pollution	>1	>1	0.611
... to income	>1	>1	>1
	Panel C: 2019 and after		
Partial R^2	0.047	0.054	0.021
Robustness value	0.111	0.127	0.0404
... to violent protest	0.415	0.678	0.0348
... to property crime	0.431	0.745	0.0457
... to pollution	>1	>1	0.201
... to income	>1	>1	>1

Note: This table describes the sensitivity of each of the significant results of Table 2 to violations of the exclusion restriction. For each result, we follow Cinelli and Hazlett (2020b) and report the partial R^2 of the instrument in the reduced form regression, as well as the robustness value $RV_{q,\alpha}$ for a significance level $\alpha = 0.10$: the results indicate that to overturn the result, an exclusion restriction violating variable Z would have to have a partial R^2 with weather of at least this value, and a partial R^2 with the outcome above this value. The following lines take different examples of possible variables through which the exclusion restriction is violated. For each of these variables, taking account a typical value of their partial correlation with the weather (Mellon, 2021), the value indicate the strength of the partial correlation with the outcome necessary to overturn the result. The partial correlations used as input are 0.0454 for violent protest, 0.0445 for property crime, 0.0131 for pollution and 0.0075 for income. These values have been computed using the `sensemakr` Stata command (Cinelli et al., 2020).

Table A4: Second stage regression varying weather controls.

Outcomes	Left-right axis				Corruption		
	Left	Center	Right	Far-right	Anti-corruption	Not anti-corruption	Abstention
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Panel A: Unpleasant weather only						
Ratio demonstrators	0.47	0.080	-0.40	-0.32**	0.72**	-0.55	-0.17
	(0.44)	(0.10)	(0.41)	(0.13)	(0.28)	(0.48)	(0.51)
AR test p-value	0.29	0.42	0.35	0.0026***	0.0039***	0.26	0.74
	Panel B: ... and prob. unp. weather in June (excl. 2011)						
Ratio demonstrators	1.06**	0.066	-0.59	-0.28**	0.86***	-0.33	-0.54
	(0.49)	(0.12)	(0.52)	(0.14)	(0.32)	(0.60)	(0.64)
AR test p-value	0.021**	0.58	0.24	0.021**	0.0016***	0.59	0.40
	Panel C: ... and squared term						
Ratio demonstrators	0.99**	0.055	-0.60	-0.28**	0.85***	-0.39	-0.46
	(0.46)	(0.12)	(0.50)	(0.13)	(0.31)	(0.59)	(0.61)
AR test p-value	0.028**	0.64	0.22	0.017**	0.0012***	0.50	0.46
	Panel D: Prob. unp. weather in June (2011 and excl. 2011)						
Ratio demonstrators	1.10**	0.073	-0.57	-0.30**	0.87***	-0.26	-0.61
	(0.50)	(0.13)	(0.54)	(0.14)	(0.33)	(0.62)	(0.64)
AR test p-value	0.021**	0.55	0.27	0.021**	0.0022***	0.67	0.34
	Panel F: All weather controls (main regression)						
Ratio demonstrators	1.03**	0.062	-0.57	-0.29**	0.86***	-0.33	-0.53
	(0.47)	(0.12)	(0.52)	(0.14)	(0.32)	(0.60)	(0.61)
AR test p-value	0.028**	0.61	0.26	0.018**	0.0018***	0.58	0.40
Mean of dependent variable	0.309	0.043	0.276	0.040	0.132	0.519	0.349
Observations	534	534	534	534	534	534	534
Election fixed effects	Y	Y	Y	Y	Y	Y	Y
Non-weather controls	Y	Y	Y	Y	Y	Y	Y

Note: Regression of First stage regression with various combination of the expected weather controls. All models include the usual city-level controls excluding weather-related controls and election year fixed effects. The sample consists of all congressional and European elections between 2011 and 2019. Panel A does not include any weather controls. Panel B controls by the expected probability of unpleasant weather on one day in June, excluding 2011. Panel C adds a square term for this probability. Panel D instead adds the probability of observing unpleasant weather in June 2011, but excluding three days before and three days after the protest. Finally, panel E presents the main specification. The corresponding first stage results are presented in Table 1. Standard errors (in parentheses) are clustered at the municipality level. AR p-values are robust to weak instruments, following (Mikusheva and Poi, 2006; Anderson and Rubin, 1949). (*10%, **5%, ***1%).

Table A5: Placebo estimate of the first stage.

	(1) Unemployment rate (proxy)
Unpleasant weather	0.010
	(0.0081)
Probability of unpleasant weather	0.0035
	(0.022)
Population (municipality)	Y
Voting intentions before 15M	Y
F statistic unpleasant weather	1.59
R^2	0.40
Observations	89

Note: Regression of the ratio of unemployed people over the population between 16 and 65 years old (a proxy variable for unemployment rate) on unpleasant weather during the June 19, 2011 demonstration. The model includes the usual city-level controls, except for unemployment rate. Robust standard errors in parentheses (*10%, **5%, ***1%).

Table A6: Placebo effects elections prior to 15M

Outcomes	Left-right axis				Corruption		
	Left	Center	Right	Far-right	Anti-corruption	Not anti-corruption	Abstention
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Panel A: 2000 congressional elections						
Ratio demonstrators	-0.42	-0.11	0.84	-0.011*	-0.35	0.75	-0.40
	(0.70)	(0.34)	(0.80)	(0.0062)	(0.22)	(0.65)	(0.72)
Mean of dependent variable	0.304	0.007	0.363	0.001	0.057	0.637	0.306
	Panel B: 2004 congressional elections						
Ratio demonstrators	-0.37	0.063	0.58	-0.014	-0.28	0.71	-0.43
	(0.75)	(0.043)	(0.73)	(0.012)	(0.20)	(0.62)	(0.63)
Mean of dependent variable	0.390	0.003	0.341	0.002	0.055	0.704	0.241
	Panel B: 2008 congressional elections						
Ratio demonstrators	0.42	-0.12	0.39	0.023	-0.18	0.92	-0.74
	(0.50)	(0.086)	(0.66)	(0.019)	(0.13)	(0.62)	(0.64)
Mean of dependent variable	0.370	0.014	0.347	0.002	0.041	0.707	0.252
City-level controls	Y	Y	Y	Y	Y	Y	Y
Observations	89	89	89	89	89	89	89

Note: Placebo estimation of effect of the ratio of demonstrators in a city, instrumented by unpleasant weather, on aggregated electoral outcomes, pooled over past elections. Column 1-4 aggregate all parties along a left-right axis. Respectively, column 1 includes all left-wing parties, column 2 centrist parties, column 3 right-wing parties (including far-right) and column 4 far-right parties. Column 5 shows results for abstention. Panel A pools together four past elections (congressional elections of 2004 and 2008, and European elections of 2004 and 2009). Panel B and C consider the 2004 and 2008 elections separately, respectively. All models include the usual city-level controls and election fixed effects. Robust standard errors in parentheses (*10%, **5%, ***1%).

Table A7: Placebo effects on attitudes prior to 15M

OUTCOMES	Corruption main worry	Political self-classification			
		Left	Right	Far-right	Don't know
	(1)	(2)	(3)	(4)	(5)
	Panel A: May 2010 to April 2011				
Ratio demonstrators	0.00730	1.163	0.260	-0.286	-1.422*
	(0.0373)	(0.784)	(0.624)	(0.175)	(0.825)
F first stage	39.67	38.93	38.93	38.93	38.93
Mean of dependent variable	0.00828	0.618	0.259	0.0295	0.123
Observations	12,317	11,128	11,128	11,128	11,128
City-level controls	Y	Y	Y	Y	Y
Individual-level controls	Y	Y	Y	Y	Y
Month fixed effects	Y	Y	Y	Y	Y

Note: Placebo estimation of effect of the ratio of demonstrators in a city, instrumented by unpleasant weather, on concerns about corruption and on self-placement on the left-right axis (1-10). The sample consists of all individuals interviewed in the monthly "CIS Barometer" between May 2010 and April 2011. Column 1 shows results for responding that corruption is the main problem of Spain. Column 2-4 show the self-placement on the 1-10 left-right axis (left, right, far-right respectively). Column 5 shows the results for the "don't know" answer to the political self-placement question. All models include the usual city-level controls, individual controls, and month fixed effects. Standard errors (in parentheses) are clustered at the province level (*10%, **5%, ***1%).

Table A8: First stage regression with logarithm of demonstrators.

	(1) Log(demonstrators)
Unpleasant weather	-1.173***
	(0.346)
Probability of unpleasant weather	1.166
	(1.132)
City-level controls	Y
F statistic unpleasant weather	11.50
R^2	0.52
Observations	89

Note: Regression of the logarithm of the number of demonstrators in the June 19 2011 demonstration on unpleasant weather during this same demonstration. The model includes the usual city-level controls. Robust standard errors in parentheses (*10%, **5%, ***1%).

Table A9: Logarithm of the number of demonstrators as instrument

Outcomes	Left-right axis				Corruption		
	Left	Center	Right	Far-right	Anti-corruption	Not anti-corruption	Abstention
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Panel A: 2011 and after						
Log(demonstrators)	0.019*	0.0011	-0.010	-0.0054*	0.016***	-0.0060	-0.010
	(0.010)	(0.0021)	(0.010)	(0.0029)	(0.0057)	(0.011)	(0.011)
Mean of dependent variable	0.306	0.042	0.289	0.034	0.123	0.536	0.341
Observations	623	623	623	623	623	623	623
	Panel B: 2014 and after						
Log(demonstrators)	0.019*	0.0020	-0.011	-0.0063*	0.018***	-0.0076	-0.010
	(0.010)	(0.0020)	(0.010)	(0.0033)	(0.0063)	(0.011)	(0.011)
Mean of dependent variable	0.309	0.043	0.276	0.040	0.132	0.519	0.349
Observations	534	534	534	534	534	534	534
	Panel C: 2019 and after						
Log(demonstrators)	0.022*	0.00021*	-0.018	-0.013**	0.0076	-0.0044	-0.0031
	(0.011)	(0.00012)	(0.013)	(0.0064)	(0.0047)	(0.012)	(0.011)
Mean of dependent variable	0.326	0.001	0.332	0.076	0.105	0.576	0.319
Observations	267	267	267	267	267	267	267
Election fixed effects	Y	Y	Y	Y	Y	Y	Y
City-level controls	Y	Y	Y	Y	Y	Y	Y

Note: Effect of the logarithm of the number of demonstrators in a city, instrumented by unpleasant weather, on aggregated electoral outcomes, pooled over elections. Column 1-4 aggregates all parties in a right-left axis. respectively column 1 includes all left-wing parties, column 2 centrist parties, column 3 right-wing parties (including far-right) and column 4 far-right parties. Outcomes of columns 5 and 6 aggregate parties in a anti-corruption not anti-corruption axis. Column 7 shows results for abstention. Panel A pools together all elections starting from 2011 (2011, 2015, 2016, April 2019 and November 2019 general elections, and 2014 and May 2019 European elections). Panel B pools only the elections starting in 2014, and panel C only the 2019 elections. All models include the usual city-level controls and election fixed effects. Standard errors (in parentheses) are clustered at the municipality level (*10%, **5%, ***1%).

Table A10: Additional controls

Outcomes	Left-right axis				Corruption		
	Left (1)	Center (2)	Right (3)	Far-right (4)	Anti-corruption (5)	Not anti-corruption (6)	Abstention (7)
	Panel A: Controls for Catalonia and Basque Country						
Ratio demonstrators	1.11**	0.049	-0.42	-0.26*	0.81***	-0.088	-0.72
	(0.52)	(0.12)	(0.52)	(0.15)	(0.31)	(0.57)	(0.59)
Mean of dependent variable	0.306	0.042	0.289	0.034	0.123	0.536	0.341
Observations	623	623	623	623	623	623	623
	Panel B: Control for GDP per capita						
Ratio demonstrators	0.81**	0.033	-0.29	-0.23**	0.62***	-0.052	-0.57
	(0.36)	(0.10)	(0.40)	(0.10)	(0.19)	(0.49)	(0.54)
Mean of dependent variable	0.306	0.042	0.289	0.034	0.123	0.536	0.341
Observations	623	623	623	623	623	623	623
Election fixed effects	Y	Y	Y	Y	Y	Y	Y
City-level controls	Y	Y	Y	Y	Y	Y	Y

Note: Effect of the ratio of demonstrators in a city, instrumented by unpleasant weather, on aggregated electoral outcomes, pooled over elections. Column 1-4 aggregates all parties in a right-left axis. respectively column 1 includes all left-wing parties, column 2 centrist parties, column 3 right-wing parties (including far-right) and column 4 far-right parties. Outcomes of columns 5 and 6 aggregate parties in a anti-corruption not anti-corruption axis. Column 7 shows results for abstention. Both panels present the effect for the subsample including all elections after 2011. Panel A shows the results when adding separate dummy variables for Catalonia and Basque Country in addition to usual controls. Panel B shows the results when adding variable for GDP per capita in the province. All models also include the usual city-level controls and election fixed effects. Standard errors (in parentheses) are clustered at the municipality level (*10%, **5%, ***1%).

Table A11: First stage regression with additional controls.

Outcomes	Ratio demonstrators	
	(1)	(2)
Unpleasant weather	-0.0228***	-0.0208***
	(0.00689)	(0.00667)
Probability of unpleasant weather	0.00862	0.0127
	(0.0252)	(0.0257)
City-level controls	Y	Y
Catalonia and Basque country	Y	
GDP per capita (province)		Y
Mean of dependent variable	0.033	0.033
F statistic for unpleasant weather	10.96	9.70
R^2	0.36	0.36
Observations	89	89

Note: First stage regression for Table A10. Regression of the ratio of demonstrators in the June 19 2011 demonstration on unpleasant weather during this same demonstration. The models includes the usual city-level controls. Additionally, column 1 includes separate dummy variables for Catalonia and Basque Country, and column 2 includes a control for the province's GDP per capita. Robust standard errors in parentheses (*10%, **5%, ***1%).

Table A12: Main results, with p-values robust to weak instruments

OUTCOMES	Left-right axis				Corruption		
	Left (1)	Center (2)	Right (3)	Far-right (4)	Anti-corruption (5)	Not anti-corruption (6)	Abstention (7)
	Panel A: 2011 and after						
Ratio demonstrators	1.03**	0.062	-0.57	-0.29**	0.86***	-0.33	-0.53
	(0.47)	(0.12)	(0.52)	(0.14)	(0.32)	(0.60)	(0.61)
AR test p-value	0.028**	0.61	0.26	0.018**	0.0018***	0.58	0.40
Mean of dependent variable	0.306	0.042	0.289	0.034	0.123	0.536	0.341
Observations	623	623	623	623	623	623	623
	Panel B: 2014 and after						
Ratio demonstrators	1.05**	0.11	-0.61	-0.35**	0.97***	-0.42	-0.55
	(0.48)	(0.12)	(0.52)	(0.16)	(0.36)	(0.63)	(0.62)
AR test p-value	0.026**	0.34	0.22	0.017**	0.0014***	0.49	0.38
Mean of dependent variable	0.309	0.043	0.276	0.040	0.132	0.519	0.349
Observations	534	534	534	534	534	534	534
	Panel C: 2019 and after						
Ratio demonstrators	1.19**	0.011*	-0.97	-0.69**	0.42*	-0.24	-0.17
	(0.54)	(0.0068)	(0.65)	(0.31)	(0.25)	(0.63)	(0.61)
AR test p-value	0.023**	0.068*	0.11	0.013**	0.091*	0.70	0.78
Mean of dependent variable	0.326	0.001	0.332	0.076	0.105	0.576	0.319
Observations	267	267	267	267	267	267	267
Election fixed effects	Y	Y	Y	Y	Y	Y	Y
City-level controls	Y	Y	Y	Y	Y	Y	Y

Note: Effect of the ratio of demonstrators in a city, instrumented by unpleasant weather, on aggregated electoral outcomes, pooled over elections. Column 1-4 aggregate all parties along a left-right axis. Respectively, column 1 includes all left-wing parties, column 2 centrist parties, column 3 right-wing parties (including far-right) and column 4 far-right parties. Outcomes of columns 5 and 6 aggregate parties along an anti-corruption / not anti-corruption axis. Column 7 shows results for abstention. Panel A pools together all elections starting from 2011 (2011, 2015, 2016, April 2019 and November 2019 general elections, and 2014 and May 2019 European elections). Panel B pools only the elections starting in 2014, and panel C only the 2019 elections. All models include city-level controls: probability of unpleasant weather in June, variance of unpleasant weather in June, probability of unpleasant weather in June 2011 excluding three days before and three days after the day of the demonstration, population, unemployment rates and the percentage of intention to vote for different parties just before the protest. All models also include election fixed effects. Standard errors (in parentheses) are clustered at the municipality level. AR p-values are robust to weak instruments, following (Mikusheva and Poi, 2006; Anderson and Rubin, 1949). (*10%, **5%, ***1%).

Table A13: Adjusting standard errors for spatial correlation

OUTCOMES	Left (1)	Center (2)	Right (3)	Far-right (4)	Anti-corruption (5)	Not anti-corruption (6)	Abstention (7)
				Panel A: 2011 and after			
Ratio demonstrators	1.03	0.062	-0.57	-0.29	0.86	-0.33	-0.53
Standard error — 50 km	(0.28)***	(0.081)	(0.38)	(0.12)**	(0.16)***	(0.42)	(0.39)
— 100 km	(0.44)**	(0.082)	(0.40)	(0.11)***	(0.18)***	(0.46)	(0.50)
— 150 km	(0.38)***	(0.083)	(0.44)	(0.11)***	(0.083)***	(0.55)	(0.52)
— 200 km	(0.38)***	(0.11)	(0.38)	(0.12)**	(0.22)***	(0.45)	(0.44)
Mean of dependent variable	0.306	0.042	0.289	0.034	0.123	0.536	0.341
Observations	623	623	623	623	623	623	623
				Panel B: 2014 and after			
Ratio demonstrators	1.05	0.11	-0.61	-0.35	0.97	-0.42	-0.55
Standard error — 50 km	(0.30)***	(0.091)	(0.39)	(0.14)**	(0.19)***	(0.44)	(0.40)
— 100 km	(0.44)**	(0.092)	(0.40)	(0.13)***	(0.20)***	(0.46)	(0.50)
— 150 km	(0.38)***	(0.093)	(0.43)	(0.12)***	(0.10)***	(0.55)	(0.51)
— 200 km	(0.40)***	(0.11)	(0.39)	(0.14)**	(0.24)***	(0.45)	(0.44)
Mean of dependent variable	0.309	0.043	0.276	0.040	0.132	0.519	0.349
Observations	534	534	534	534	534	534	534
				Panel C: 2019 and after			
Ratio demonstrators	1.19	0.011	-0.97	-0.69	0.42	-0.24	-0.17
Standard error — 50 km	(0.41)***	(0.0065)*	(0.54)*	(0.24)***	(0.20)**	(0.54)	(0.48)
— 100 km	(0.53)**	(0.0083)	(0.54)*	(0.22)***	(0.22)*	(0.58)	(0.53)
— 150 km	(0.46)***	(0.0087)	(0.56)*	(0.21)***	(0.20)**	(0.67)	(0.53)
— 200 km	(0.51)**	(0.0085)	(0.56)*	(0.24)***	(0.27)	(0.58)	(0.46)
Mean of dependent variable	0.309	0.043	0.276	0.040	0.132	0.519	0.349
Observations	267	267	267	267	267	267	267
Election fixed effects	Y	Y	Y	Y	Y	Y	Y
City-level controls	Y	Y	Y	Y	Y	Y	Y

Note: Effect of the ratio of demonstrators in a city, instrumented by unpleasant weather, on aggregated electoral outcomes, pooled over elections. Column 1-4 aggregates all parties in a right-left axis. respectively column 1 includes all left-wing parties, column 2 centrist parties, column 3 right-wing parties (including far-right) and column 4 far-right parties. Outcomes of columns 5 and 6 aggregate parties in a anti-corruption not anti-corruption axis. Column 7 shows results for abstention. Both panels present the effect for the subsample including all elections after 2011. Panel A pools together all elections starting from 2011 (2011, 2015, 2016, April 2019 and November 2019 general elections, and 2014 and May 2019 European elections). Panel B pools only the elections starting in 2014, and panel C only the 2019 elections. All models include the usual city-level controls and election fixed effects. Standard errors (in parentheses) are computed taking into account spatial correlation, for multiple threshold distances after which the errors are considered to be uncorrelated: 50 km, 100 km, 150 km and 200 km. The starts next to standard errors correspond to the significance level for this distance (*10%, **5%, ***1%).

Table A14: First stage regression, clustered by province.

	Ratio demonstrators (1)
Unpleasant weather	-0.0214***
	(0.00708)
Probability of unpleasant weather	0.00804
	(0.0265)
City-level controls	Y
Mean of dependent variable	0.033
F statistic for unpleasant weather	9.133
R^2	0.35
Observations	89

Note: Regression of the ratio of demonstrators in the June 19 demonstration over the population of the city on unpleasant weather during this same demonstration. The model includes the usual city-level controls, as well as a control for the probability of a day in May or July 2011 having pleasant weather. Standard errors (in parentheses) are clustered at the province level (*10%, **5%, ***1%).

Table A15: Clustering by province

OUTCOMES	Left-right axis				Corruption		
	Left (1)	Center (2)	Right (3)	Far-right (4)	Anti-corruption (5)	Not anti-corruption (6)	Abstention (7)
				Panel A: 2011 and after			
Ratio demonstrators	1.03**	0.062	-0.57	-0.29**	0.86***	-0.33	-0.53
	(0.50)	(0.12)	(0.58)	(0.14)	(0.30)	(0.69)	(0.70)
AR test p-value	0.035**	0.60	0.32	0.027**	0.0017***	0.64	0.45
Mean of dependent variable	0.306	0.042	0.289	0.034	0.123	0.536	0.341
Observations	623	623	623	623	623	623	623
				Panel B: 2014 and after			
Ratio demonstrators	1.05**	0.11	-0.61	-0.35**	0.97***	-0.42	-0.55
	(0.50)	(0.12)	(0.58)	(0.16)	(0.35)	(0.72)	(0.70)
AR test p-value	0.035**	0.33	0.28	0.026**	0.0015***	0.56	0.43
Mean of dependent variable	0.309	0.043	0.276	0.040	0.132	0.519	0.349
Observations	534	534	534	534	534	534	534
				Panel C: 2019 and after			
Ratio demonstrators	1.19**	0.011	-0.97	-0.69**	0.42*	-0.24	-0.17
	(0.56)	(0.0076)	(0.70)	(0.31)	(0.24)	(0.73)	(0.70)
AR test p-value	0.027**	0.12	0.16	0.020**	0.081*	0.74	0.80
Mean of dependent variable	0.326	0.001	0.332	0.076	0.105	0.576	0.319
Observations	267	267	267	267	267	267	267
Election fixed effects	Y	Y	Y	Y	Y	Y	Y
City-level controls	Y	Y	Y	Y	Y	Y	Y

Note: Effect of the ratio of demonstrators in a city, instrumented by unpleasant weather, on aggregated electoral outcomes, pooled over elections. Column 1-4 aggregate all parties along a left-right axis. Respectively, column 1 includes all left-wing parties, column 2 centrist parties, column 3 right-wing parties (including far-right) and column 4 far-right parties. Outcomes of columns 5 and 6 aggregate parties along an anti-corruption / not anti-corruption axis. Column 7 shows results for abstention. Panel A pools together all elections starting from 2011 (2011, 2015, 2016, April 2019 and November 2019 general elections, and 2014 and May 2019 European elections). Panel B pools only the elections starting in 2014, and panel C only the 2019 elections. All models include city-level controls and election fixed effects. Standard errors (in parentheses) are clustered at the province level. AR p-values are robust to weak instruments, following (Mikusheva and Poi, 2006; Anderson and Rubin, 1949). (*10%, **5%, ***1%).

Table A16: Romano-Wolf correction for multiple hypothesis testing

OUTCOMES	Left-right axis				Corruption		
	Left (1)	Center (2)	Right (3)	Far-right (4)	Anti-corruption (5)	Not anti-corruption (6)	Abstention (7)
	Panel A: 2011 and after						
Ratio demonstrators	1.03*	0.062	-0.57	-0.29*	0.86*	-0.33	-0.53
	(0.47)	(0.12)	(0.52)	(0.14)	(0.32)	(0.60)	(0.61)
Corrected p-value	0.078	0.727	0.459	0.075	0.073	0.960	0.600
Mean of dependent variable	0.306	0.042	0.289	0.034	0.123	0.536	0.341
Observations	623	623	623	623	623	623	623
	Panel B: 2014 and after						
Ratio demonstrators	1.05*	0.11	-0.61	-0.35*	0.97*	-0.42	-0.55
	(0.48)	(0.12)	(0.52)	(0.16)	(0.36)	(0.63)	(0.62)
Corrected p-value	0.079	0.460	0.418	0.078	0.079	0.913	0.460
Mean of dependent variable	0.309	0.043	0.276	0.040	0.132	0.519	0.349
Observations	534	534	534	534	534	534	534
	Panel C: 2019 and after						
Ratio demonstrators	1.19*	0.011	-0.97	-0.69*	0.42*	-0.24	-0.17
	(0.54)	(0.0068)	(0.65)	(0.31)	(0.25)	(0.63)	(0.61)
Corrected p-value	0.063	0.182	0.182	0.055	0.122	0.805	0.805
Mean of dependent variable	0.326	0.001	0.332	0.076	0.105	0.576	0.319
Observations	267	267	267	267	267	267	267
Election fixed effects	Y	Y	Y	Y	Y	Y	Y
City-level controls	Y	Y	Y	Y	Y	Y	Y

Note: Effect of the ratio of demonstrators in a city, instrumented by unpleasant weather, on aggregated electoral outcomes, pooled over elections. Column 1-4 aggregates all parties in a right-left axis. respectively column 1 includes all left-wing parties, column 2 centrist parties, column 3 right-wing parties (including far-right) and column 4 far-right parties. Outcomes of columns 5 and 6 aggregate parties in a anti-corruption not anti-corruption axis. Column 7 shows results for abstention. Panel A pools together all elections starting from 2011 (2011, 2015, 2016, April 2019 and November 2019 general elections, and 2014 and May 2019 European elections). Panel B pools only the elections starting in 2014, and panel C only the 2019 elections. All models include the usual city-level controls and election fixed effects. Standard errors (in parentheses) are clustered at the province level. Corrected p-values are obtained using the Romano-Wolf procedure to correct for multiple hypothesis testing (Romano and Wolf, 2005) for each individual row (*10%, **5%, ***1%).

B Choice of instrument

In this section I explain the rationale behind the choice of the weather conditions that constitute "pleasant weather", and the controls for expected weather used in the identification. The main identification strategy uses a dummy variable for unpleasant weather, where weather is considered unpleasant if the temperature is above 30°C, or if there is any rain. I also use three controls for the expected weather: I control for regional weather patterns by including both the probability of unpleasant weather on a day in June but not in 2011 and its square, and the probability of unpleasant weather on a day in June 2011, excluding the three days before and after the demonstration. I will discuss the effect of all these components on the identification.

B.1 Choice of temperature and rain threshold

I consider any day with rain to be unpleasant: the literature has consistently found a negative relationship between the likelihood and the magnitude of a protest and the presence of rain (Madestam et al., 2013; Zhang, 2016; Larreboure and González, 2021). Only two cities (out of 89) are classified as having unpleasant weather because of rain, and one of them would also be classified as unpleasant due to the temperature: only Girona would switch from pleasant to unpleasant if we were to exclude rain. I show in Appendix A that most results are not sensitive to the exclusion of one city.

The choice of the threshold in temperature is particularly important because it is the factor from which the majority of the variation in the classification between pleasant and unpleasant comes from. It has previously been documented in the literature (Zhang, 2016) that the likelihood of observing a protest with respect to temperature follows an inverted U shape peaking at around 20ºC and decreasing afterwards.

Figure B1a shows the coefficients of the first stage using different temperature thresholds with their respective confidence interval (scale on the left-hand side Y axis) and the number of cities that will qualify as having unpleasant weather if each threshold would have been used (scale on the right-hand side Y axis). Figure B1b shows the F statistic of the measure of unpleasant weather for each of these temperatures. Only three temperature thresholds are significant for explaining protest participation: 30°C, 31°C and 32°C. 30°C is the temperature that gives the largest (in absolute value) significant coefficient. This is also the best first stage in terms of the F statistic: a temperature threshold of 30°C is the only one producing a F statistic above the standard threshold of 10 for strong instruments. Moreover, using it as a threshold produces greater variability among cities with unpleasant weather as it is the one that excludes the least. Under this specification, 28 out of 89 cities are considered to have had unpleasant weather during the day of the march.

B.2 Choice of expected weather controls

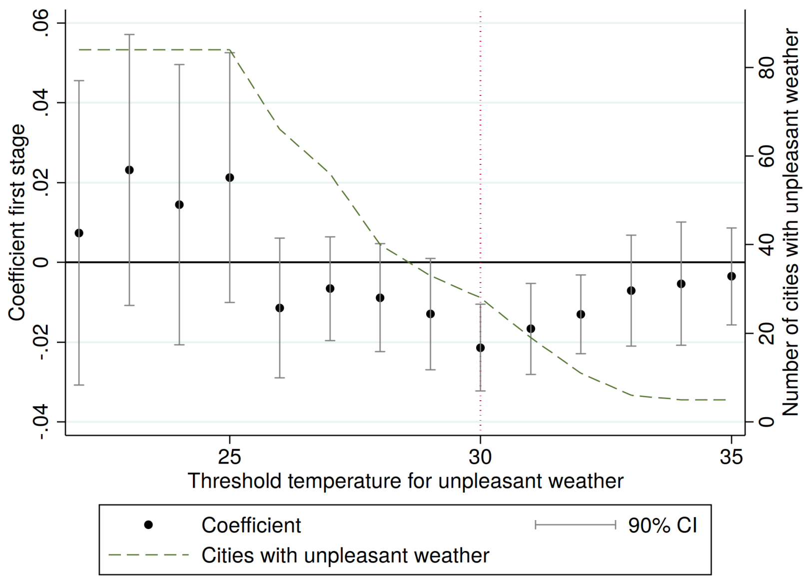
The regression uses three controls for expected weather. These controls play different roles. The inclusion of the probability of unpleasant weather as a control is particularly important. Indeed cities that experience unpleasant weather more frequently (for example, inner regions of the South of Spain), are likely to have different protest behavior and different electoral results than cities with lower probability of unpleasant weather (for example, areas in the Atlantic coast). The inclusion of the probability of unpleasant weather means the instrument can be interpreted as a weather shock, i.e. the deviation from the probability of experiencing an unpleasant weather on June 19, 2011 with respect to an average month of June in the city. The inclusion of a squared term produces a flexible functional form able to capture non linear effects of weather conditions and increase precision.

The inclusion of the probability of unpleasant weather on June 2011 captures a different phenomenon than the probability of unpleasant weather in other years. As weather is temporally correlated, experiencing deviation from mean weather conditions on the day of the demonstration is likely correlated to weather deviation in the same direction on the surrounding days and weeks. The exclusion restriction could be violated if weather conditions in surrounding days affect other characteristics that can impact political preferences. Controlling for these weather shocks alleviates this concern and get us closer to being able to interpret the instrument as capturing only short-run effects of weather. I exclude the six days (three days before and three days after) surrounding the day of the demonstration to leave enough variation in weather conditions of the day of the demonstration.

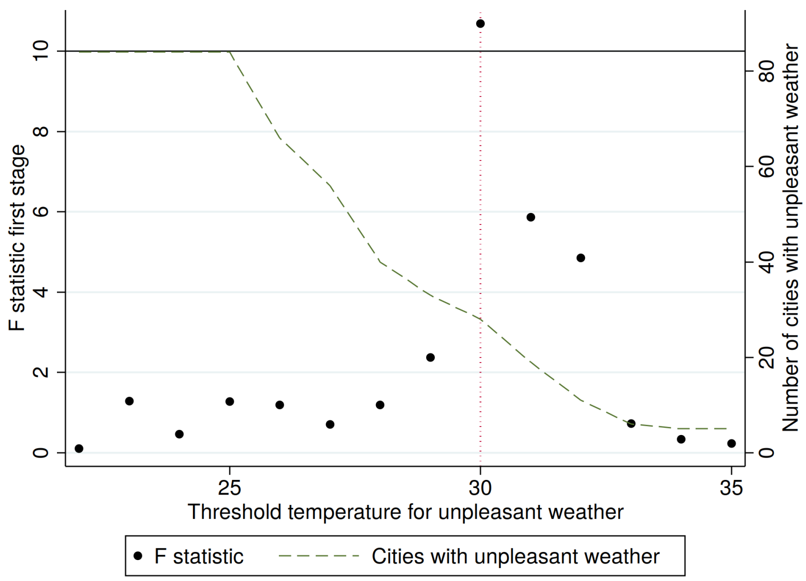
Effect of removing some controls In Table 1, I remove some controls: column 1 is the first stage without expected weather controls, column 2 uses only the expected weather excluding 2011, column 3 adds the square term, column 4 uses the expected weather excluding 2011 and the expected weather in 2011, and column 6 presents the first stage for the main regression. Neither the coefficient nor the significance of unpleasant weather vary much, and the coefficient of the controls stay relatively small and are non-significant expect for the probability of unpleasant weather in June 2011. This indicates (perhaps surprisingly) that these controls do not majorly influence the first stage. A possible interpretation is that the weather conditions on June 19th, 2011 were very unusual and even in regions used to hot weather, people were nonetheless affected by the heat. I present and discuss in Table A4 and in Appendix A the second stage results for the main regression.

Figure B1: First stage regression when varying the threshold temperature for unpleasant weather

(a) Coefficient and confidence interval



(b) Coefficient and confidence interval



Note: Plot of the estimated coefficient of the first stage regression (additional percentage points of population attending the June 19 protest) and 90% confidence intervals with the threshold temperature for unpleasant weather set to the given temperature, and plot of the F statistic for unpleasant weather. The scale for the estimate and confidence interval is on the left-hand side of the graph. The red dotted line indicates the threshold temperature I use. The green dashed line gives the number of cities that are considered to have unpleasant weather if the threshold is set to the given temperature (on the right-hand-side scale). Standard errors are robust to heteroscedasticity.

C Supplementary results and interpretations

In this Appendix, I present: i) additional elements to interpret the magnitudes of the estimates of the main results in terms of persuasion and ii) a more detailed analysis of the main electoral results splitting the analysis both year by year and by the main political parties.

C.1 Persuasion

Another way to interpret the magnitude of the estimated results is in terms of persuasion (DellaVigna and Gentzkow, 2010): the percentage of the voters that would not have voted for a particular party in the absence of 15M that are convinced to vote for that party as a consequence of the movement. The persuasion rate is defined as follows:

$$\text{Persuasion rate} = \frac{\text{Population convinced by the intervention}}{\text{Unconvinced population without the intervention}}$$

Table C1 shows the estimated persuasion rates for three different interventions: a marginal increase of the number of demonstrators by 1 percentage points of the population of the city; a shift to having pleasant weather everywhere compared to the actual weather, and a shift to having the actual weather instead of having unpleasant weather everywhere. The persuasion rates higher for the far-right compared to its coefficient in the regression: this reflects the fact that the population that can be convinced in that case is the population that votes for the far-right, which is smaller than the population used in the other rates (for example the population that does not vote for anti-corruption options). The persuasion rates are generally high: for example, switching all cities to pleasant weather would remove 6% of the vote for the far-right in elections in 2019, which is a significant effect for a single demonstration.

C.2 Disagregated results

In this subsection I show and briefly interpret the results when disagregating: i) the sub-samples pooling together different election, in each congressional election and ii) the pooled political option in two axis in the major individual political parties.

C.2.1 Each congressional election separately

Table C2 presents the results for each congressional election that occurred after the movement up to 2019 in a different panel. Overall, results show a general leftward shift in political preferences. All panels show an increased vote for left-wing parties due to higher protest attendance in the city. In the 2011 elections (panel A), when the new political option that is specially linked to the 15M movement (Podemos) was not yet created, we see an increase in protest vote,footnoteProtest vote is defined as casting a blank or null ballot. showing a willingness to participate in the democratic institutions (if this was not the case we would probably observe an increase in abstention) but clearly expressing discontent. The effect for the far-right is mainly present for elections where far-right parties were stronger (panel D and E). Finally, the effects for anti-corruption parties is mainly present between 2014 and the first election of 2019. Before 2014, neither Podemos nor Ciudadanos (the two main anti-corruption parties) were credible political options and after the middle of 2019, vote for both of these parties had radically decreased. In addition, corruption being a more and more prominent concern in Spain, more and more parties have started promoting measures to fight it. On the one hand, this reduces the number of major corruption scandals that are mentioned in the media (the last major scandal was in 2018). On the other hand, parties that do not talk about corruption have become less frequent, reducing then the variation that is necessary to compute precise estimates.

C.2.2 Disaggregated options

Table C3 presents the results for each main political party separately. Overall the results show that there is not a single political party that can explain the overall persistent shift to the left. Panel B shows a strong effect for Podemos showing that cities with higher protest attendance increase the vote for Podemos after its creation. However this effect fades over time and is not statistically significant for the elections of 2019. The reduction in the vote for Vox (the far-right party of Spain) is present for all sub-samples but more so for elections where it was stronger suggesting a more intense anti-Vox campaign in cities with higher level of attendance to 15M demonstrations.

Table C1: Persuasion effects of different treatments on voting outcomes.

	1 pp. increase in demonstrators (1)	From unpleasant to actual weather (2)	From actual weather to pleasant weather (3)
		2011 and after	
Voting for the left	1.48%	1.98%	1.19%
Not voting for the far-right	8.41%	9.95%	6.50%
Voting for anti-corruption options	0.98%	1.29%	0.77%
		2014 and after	
Voting for the left	1.52%	2.08%	1.25%
Not voting for the far-right	8.73%	10.1%	6.60%
Voting for anti-corruption options	1.12%	1.52%	0.91%
		2019 and after	
Voting for the left	1.76%	2.32%	1.40%
Not voting for the far-right	9.11%	10.9%	7.23%
Voting for anti-corruption options	0.47%	0.63%	0.37%

Note: Columns show different significant electoral effects and rows different counterfactual treatments, for all significant effects in Table 2. The persuasion effect is the effect on the percentage of the voters that will not have voted for a particular party in the absence of 15M that are convinced to vote for that party as a consequence of the movement. This interpretation of the results takes into account that only the people that do not already vote for a party can be persuaded to vote for it, and only the people that vote for a party can be persuaded to stop voting for it. Column (1) shows the effect of a 1 pp. increase in demonstrators in all cities, column (2) shows the effect of the actual weather compared to the counterfactual where all cities experienced unpleasant weather, and column (3) shows the effect of the counterfactual where all cities experienced pleasant weather compared with the actual weather on June 19 (both columns 2 and 3 are evaluated from the reduced form estimate).

Table C2: Electoral effects of protest attendance on each congressional election

Outcomes	Left-right axis				Corruption		
	Left (1)	Center (2)	Right (3)	Far-right (4)	Anti-corruption (5)	Not anti-corruption (6)	Abstention (7)
	Panel A: 2011 congressional election						
Ratio demonstrators	0.87*	-0.22	-0.32	0.019	0.18	0.20	-0.38
	(0.52)	(0.19)	(0.58)	(0.016)	(0.20)	(0.57)	(0.64)
Mean of dependent variable	0.284	0.037	0.366	0.001	0.071	0.638	0.291
	Panel B: 2015 congressional election						
Ratio demonstrators	0.99**	0.25	-0.47	-0.032*	1.79***	-0.99	-0.79
	(0.49)	(0.31)	(0.48)	(0.016)	(0.63)	(0.76)	(0.70)
Mean of dependent variable	0.348	0.111	0.245	0.002	0.188	0.536	0.276
	Panel C: 2016 congressional election						
Ratio demonstrators	1.09**	0.10	-0.34	-0.0064	1.46***	-0.61	-0.85
	(0.48)	(0.23)	(0.50)	(0.013)	(0.53)	(0.67)	(0.63)
Mean of dependent variable	0.312	0.096	0.265	0.002	0.154	0.540	0.306
	Panel D: April 2019 congressional election						
Ratio demonstrators	1.38**	0.022**	-1.10	-0.73**	0.52*	-0.21	-0.32
	(0.56)	(0.0092)	(0.72)	(0.34)	(0.30)	(0.68)	(0.65)
Mean of dependent variable	0.351	0.000	0.370	0.080	0.121	0.624	0.254
	Panel E: November 2019 congressional election						
Ratio demonstrators	1.40**	0.015**	-0.97	-1.00**	0.20	0.16	-0.36
	(0.57)	(0.0062)	(0.69)	(0.43)	(0.29)	(0.67)	(0.66)
Mean of dependent variable	0.321	0.000	0.337	0.106	0.109	0.573	0.317
Observations	89	89	89	89	89	89	89
City-level controls	Y	Y	Y	Y	Y	Y	Y

Note: Effect of the ratio of demonstrators in a city, instrumented by unpleasant weather, on aggregated electoral outcomes. Column 1-4 aggregate all parties along a left-right axis. Respectively, column 1 includes all left-wing parties, column 2 centrist parties, column 3 right-wing parties (including far-right) and column 4 far-right parties. Outcomes of columns 5 and 6 aggregate parties along an anti-corruption / not anti-corruption axis. Column 7 shows results for abstention. Panel A present the estimates for 2011 election, panel B for 2015, panel C for 2016, panel D for the the elections of April 2019 and panel E for elections of November 2019. All models include the usual city-level controls. Robust standard errors in parentheses (*10%, **5%, ***1%).

Table C3: Electoral effects for each main party of protest attendance

Outcomes	Podemos	United Left	PSOE	Ciudadanos	PP	Vox	Protest vote	Abstention
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A: 2011 and after				
Ratio demonstrators		0.094	-0.15		0.28	-0.34**	0.016	-0.53
		(0.12)	(0.38)		(0.50)	(0.16)	(0.036)	(0.61)
Mean of dependent variable		0.039	0.177		0.198	0.040	0.015	0.341
Observations		267	623		623	623	623	623
				Panel B: 2014 and after				
Ratio demonstrators	0.80**	0.079	-0.19	-0.063	0.35	-0.34**	-0.00069	-0.55
	(0.32)	(0.11)	(0.37)	(0.20)	(0.47)	(0.16)	(0.037)	(0.62)
Mean of dependent variable	0.095	0.035	0.173	0.079	0.174	0.040	0.014	0.349
Observations	534	178	534	534	534	534	534	534
				Panel C: 2019 and after				
Ratio demonstrators	0.36		0.19	-0.25	0.60	-0.69**	-0.032	-0.17
	(0.24)		(0.39)	(0.25)	(0.43)	(0.31)	(0.045)	(0.61)
Mean of dependent variable	0.082		0.208	0.087	0.147	0.076	0.014	0.319
Observations	267		267	267	267	267	267	267
Election fixed effects	Y	Y	Y	Y	Y	Y	Y	Y
City-level controls	Y	Y	Y	Y	Y	Y	Y	Y

Note: Effect of the ratio of demonstrators in a city, instrumented by unpleasant weather, on major political parties, pooled over elections. Each column presents a the estimates for the political party at the header. Column 7 shows results for protest vote (null or blank vote) and column 8 for abstention. Panel A pools together all elections starting from 2011 (2011, 2015, 2016, April 2019 and November 2019 general elections, and 2014 and May 2019 European elections). Panel B pools only the elections starting in 2014, and panel C only the 2019 elections. All models include the usual city-level controls and election fixed effects. Standard errors (in parentheses) are clustered at the municipality level (*10%, **5%, ***1%).

D Data and data treatment

D.1 Aggregation of different data sources for protest data

I obtained the number of people in each demonstration from three different sources: 1) from the online website and Facebook page of “Democracia Real YA” –Real Democracy NOW–, a group that was very close to the organizers of the rally; 2) from a compilation available online in the web page “Toma la Plaza” –Take the Square– which is also related to the organizers; 3) from my own compilation of different local and national press sources. Each of these sources reports, in turn, data coming from different primary sources: press, organizers and police. Not every city with a demonstration had data from every source so an average should be computed. However, it is plausible that sources –particularly those coming from the police and the organizers– differ greatly, so computing a simple average would create bias. In order to address this issue, I proceed as follows: first I group the data coming from the same primary source (this is: organizers, press or police). Then I regress the data coming from press and police on the data coming from organizers. Each coefficient gives me the factor by which I need to multiply the data coming from police and the press if I want it to be at the same scale as the data coming from the organizers.[39] Once I have these results, I compute the estimated number of demonstrators that all primary sources would have reported if they were organizers. Doing that allows me to have all the data at the same scale. Finally, I compute a simple average of all the adjusted numbers of demonstrators to obtain a single measure of the number of demonstrators in each city during the demonstration of June 19. In Spain, 90 different cities had a demonstration on June 19. For missing data concerns, I exclude the demonstration held in Ceuta –a city that only has a land border to Morocco and for which other data needed for the analysis is not available –.

D.2 Aggregation of different parties in different categories

I aggregate parties in different categories along two axes: the left-right spectrum and anti-corruption parties. For the left-right spectrum, I classified the parties running in each election as in left, center, right and far-right. I count all far-right parties as right wing. The classification is based on public information about the parties, including their classification as indicated on Wikipedia. Some parties do not fit in the left-right spectrum, often because they are based around a single issue that does not clearly map on the left-right spectrum. For instance, candidates from "Escaños en Blanco" ("Blank Seats") promise that, if elected, they will leave their seat in the Parliament unoccupied.

For anti-corruption, I group all parties that highlight fighting against corruption as

[39]The results show that data coming from the organizers are, on average, 20% greater than the data coming from the press and 160% greater than the data coming from the police, confirming the previous concern.

part of their platform. The major parties that are included are United Left, Podemos and Ciudadanos until 2015. I also add votes that clearly indicate dissatisfaction with the political system: this includes blank and null votes, and votes for parties such as Escaños en Blanco. The non-anticorruption vote is all vote that is not for an anti-corruption options, excluding abstention.

D.3 Sentiment analysis using roBERTa

In Table 7, I analyze the number of tweets expressing a positive sentiment containing words connected to the 15M movement. The sentiment of tweets has been evaluated using a state-of-the art deep learning language model. More precisely, I use a multilingual XLM-roBERTa model specifically tuned for evaluating tweets (Barbieri et al., 2021).[40]

The language model has been applied to the text of the tweet for original tweets, or to the text of the retweeted message if it is a retweet. Before evaluation, the text of the tweet has been processed to replace mentions of users and links by special tokens. The classifier returns a score for each category "positive", "neutral" and "negative": I consider a tweet as belonging to the category with the highest score. In the sample, 48% of the tweets are evaluated as neutral, 39% as negative and 13% as positive.

Table D1: Most frequent words in Spanish

de	la	que	y	en	el	a	los	se	del
las	un	no	una	por	con	es	para	su	lo
como	o	al	mas	sus	ha	este	ser	pero	sobre
son	si	me	esta	entre	le	puede	muy	nos	todo
sin	cuando	todos	tambien	ya	han	desde	porque	dos	vida
mi	e	años	tiene	ni	cada	mismo	fue	parte	mundo
otros	hasta	era	hay	forma	donde	sino	solo	tiempo	uno
bien	hacer	otro	trabajo	hace	otra	vez	pueden	estos	tan
ese	caso	sido	personas	yo	gran	te	eso	siempre	sea
contra	mucho	algo	otras	todas	habia	ellos	esa	debe	les

List of the 100 most frequent words in Spanish used to search for tweets. Since Twitter search is not sensitive to accents, accents have been removed and identical words have been replaced by new words. The list of words are extracted from the "Internet" corpus of Sharoff (2006).

[40] The exact model used is available online.

Table D2: Words and hashtags related to 15M

#15M	15M	#Indignados	#SpanishRevolution
#19j	#nolesvotes	#19jmani	#democraciarealya
#dry	#globalrevolution	#marchaindignada	#worldrevolution
#soy15m	#europeanrevolution	#stopretallades	#15o
#democraciareal	#indignaos	#nonosrepresentan	#quieroserpersona
#yeswecamp	#indignats		

List of words and hashtags associated with the 15M movement and the 19th of June 2011 demonstration that were used to find related tweets.

Table D3: Principal component analysis of Twitter activity after the 19th of June 2011

(a) Correlation between measures

	Log(Tweets)	Log(Users)	Log(New users)
Log(Tweets)	1		
Log(Users)	0.990	1	
Log(New users)	0.839	0.857	1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(b) Principal components

	Eigenvalue	Difference	Proportion	Cumulative
PC1	2.792962	2.59515	0.9310	0.9310
PC2	.1978119	.1885854	0.0659	0.9969
PC3	.0092265	.	0.0031	1.0000

(c) Factor loadings

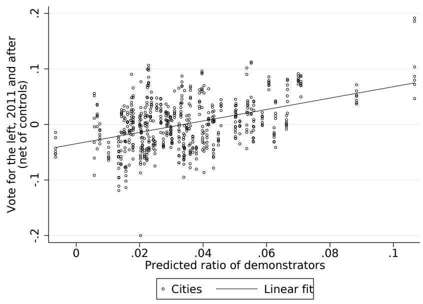
	PC1	PC2	PC3
Log(Tweets)	.5859092	-.4317025	.6858158
Log(Users)	.5894387	-.3537419	-.7262429
Log(New users)	.5561227	.8297588	.0472015

Note: Principal component analysis of the Twitter activity measures after the 19th of June 2011. The first table reports the correlation between the Twitter activity measures. The second table reports the eigenvalues of the three principal components. The third table reports the loading of the different components.

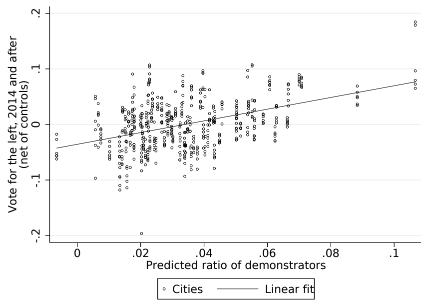
E Additional tables and figures

Figure E1: Scatter plots of the second stage (net of controls) for significant outcomes.

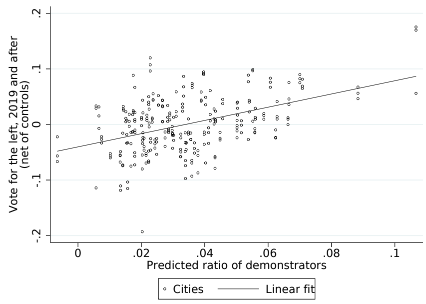
(a) Vote for the left, 2011 and after



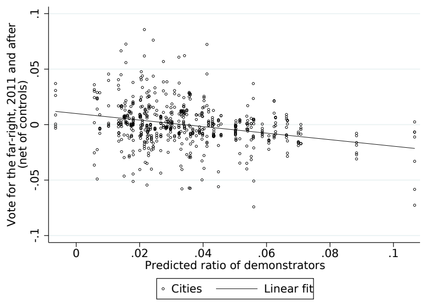
(b) 2014 and after



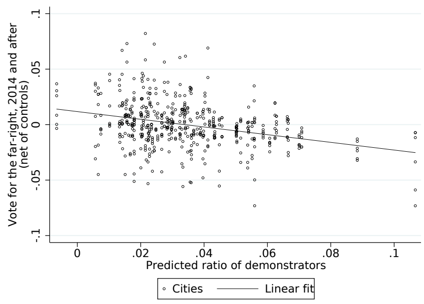
(c) 2019 and after



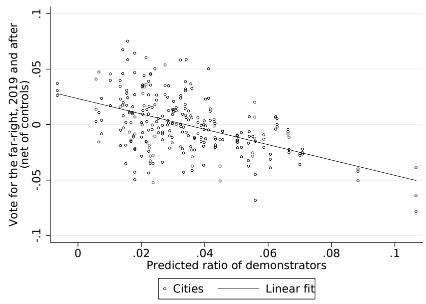
(d) Vote for far right, 2011 and after



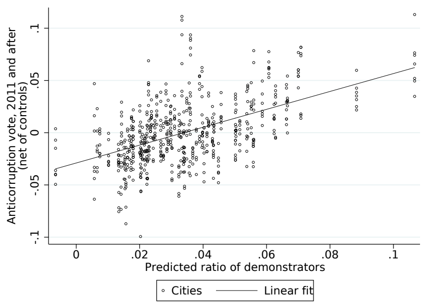
(e) Vote for far right, 2014 and after



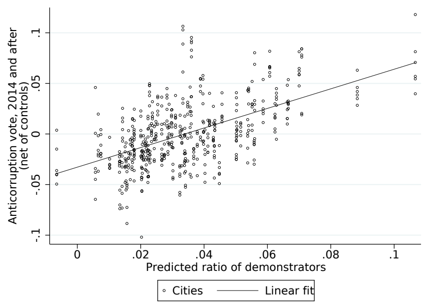
(f) Vote for far right, 2010 and after



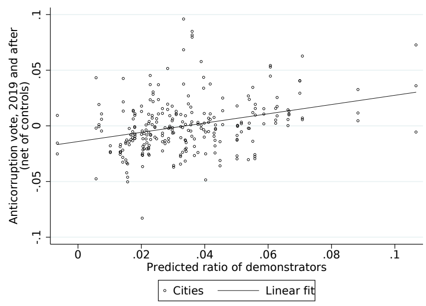
(g) Anti-corruption vote, 2011 and after



(h) Anti-corruption vote, 2014 and after



(i) Anti-corruption vote, 2019 and after



Note: Each sub-graph corresponds to one of the statistically significant outcomes of Table 2. The scatter plots show, for each city, on the Y axis the vote share of the different electoral outcomes that show statistical significance against the ratio of demonstrators predicted from the first stage regression, on the X axis. The vote shares are presented net of controls, i.e. they have been adjusted, following the second stage model, to be at the value that would be if all controls were equal to their means, preserving the residuals (in other words, they include only the effect of the predicted ratio of demonstrators and the residuals). The line shows the linear fit for the scatter plot.

Table E1: Descriptive statistics for main variables.

Variable	Obs.	Mean	St. dev.	Median	Min	Max
	City characteristics					
Ratio demonstrators	89	0.034	0.032	0.025	0.001	0.179
Unpleasant weather	89	0.315	0.467	0.000	0.000	1.000
Probability of unpleasant weather	89	0.569	0.127	0.576	0.200	0.847
Unpleasant weather in June 2011	89	0.556	0.143	0.217	0.913	
Population (municipality)	89	200068	395158	95207	9120	3265038
Proxy for unemployment	89	0.146	0.034	0.136	0.081	0.253
	Controls: Voting intention before 15M					
Traditional left	89	0.215	0.055	0.216	0.115	0.383
Traditional right	89	0.283	0.091	0.291	0.030	0.500
Traditional far-left	89	0.031	0.015	0.029	0.000	0.060
Nationalist right	89	0.014	0.044	0.000	0.000	0.192
Minor parties	89	0.059	0.038	0.052	0.000	0.248
Protest vote	89	0.055	0.027	0.054	0.000	0.115
Abstention	89	0.140	0.049	0.126	0.040	0.260
No answer	89	0.203	0.063	0.195	0.049	0.378
	Electoral outcomes (2011 and after)					
Left	623	0.306	0.065	0.308	0.136	0.514
Center	623	0.042	0.046	0.029	0.000	0.183
Right	623	0.289	0.087	0.292	0.069	0.516
Far-right	623	.034	.046	0.008	0.000	0.213
Anti-corruption	623	0.121	0.049	0.114	0.009	0.316
Non anti-corruption	623	0.539	0.115	0.568	0.184	0.754
Abstention	623	0.341	0.109	0.306	0.190	0.763
	Individual characteristics (post-15M)					
Age	288,776	49.5	17.9	48	18	99
Gender: female	288,774	0.513	0.500	1.000	0.000	1.000
Higher studies	288,776	0.278	0.448	0.000	0.000	1.000
Unemployed	288,776	0.189	0.391	0.000	0.000	1.000
Retired	288,776	0.257	0.437	0.000	0.000	1.000
	Individual outcomes (post-15M)					
Concern about corruption	287,823	0.120	0.326	0.000	0.000	1.000
Political scale: left	260,745	0.641	0.480	1.000	0.000	1.000
Political scale: right	260,745	0.243	0.429	0.000	0.000	1.000
Political scale: far right	260,745	0.030	0.170	0.000	0.000	1.000
Political scale: don't know	260,745	0.116	0.320	0.000	0.000	1.000

Table E2: Percentage of people that link each party running in the 2014 European election with the 15M movement

Party	**Percentage of people linking party to 15-M**
1. Partido X	62.1
2. Podemos	55.6
3. Primavera Europea	20.9
4. Pirata	16.2
5. Escaños en Blanco (Blank vote)	15.8
6. Izquierda Plural (Traditional far-left)	15.7
7. Movimiento RED	12.6
8. Recortes Cero	12.13
9. Partido Animalista	4.0
10. Los pueblos deciden	3.1
11. Ciudadanos	3.0
12. L'esquerra pel dret a decidir	1.2
13. PSOE (Traditional left)	0.67
14. Vox	0.7
15. UPyD	0.5
16. PP (Traditional right)	0.3
Foro Asturianas	0.3
18. Coalición por Europa (Nationalist right)	0.1

Note: Data from survey with 1208 respondents. Data source: Redes, Movimientos y Tecnopolítica (2014).

Table E3: OLS and reduced form regression

OUTCOMES	Left-right axis				Corruption		
	Left (1)	Center (2)	Right (3)	Far-right (4)	Anti-corruption (5)	Not anti-corruption (6)	Abstention (7)
	Panel A: 2011 and after – OLS						
Ratio demonstrators	0.48***	-0.047	-0.24**	-0.10***	0.32***	-0.14	-0.18
	(0.10)	(0.029)	(0.11)	(0.027)	(0.075)	(0.15)	(0.14)
Mean of dependent variable	0.306	0.042	0.289	0.034	0.123	0.536	0.341
Observations	623	623	623	623	623	623	623
	Panel B: 2011 and after – Reduced form						
Pleasant weather	0.022**	0.0013	-0.012	-0.0063**	0.018***	-0.0070	-0.011
	(0.010)	(0.0026)	(0.011)	(0.0026)	(0.0053)	(0.013)	(0.014)
Mean of dependent variable	0.306	0.042	0.289	0.034	0.120	0.539	0.341
Observations	623	623	623	623	623	623	623
	Panel C: 2014 and after – OLS						
Ratio demonstrators	0.49***	-0.037	-0.25**	-0.11***	0.34***	-0.15	-0.20
	(0.10)	(0.029)	(0.12)	(0.032)	(0.082)	(0.15)	(0.15)
Mean of dependent variable	0.309	0.043	0.276	0.040	0.129	0.522	0.349
Observations	534	534	534	534	534	534	534
	Panel D: 2014 and after – Reduced form						
Pleasant weather	0.023**	0.0023	-0.013	-0.0074**	0.021***	-0.0089	-0.012
	(0.010)	(0.0024)	(0.011)	(0.0030)	(0.0058)	(0.013)	(0.014)
Mean of dependent variable	0.309	0.043	0.276	0.040	0.132	0.519	0.349
Observations	534	534	534	534	534	534	534
	Panel E: 2019 and after – OLS						
Ratio demonstrators	0.52***	-0.0015	-0.39***	-0.22***	0.26***	-0.15	-0.11
	(0.12)	(0.0023)	(0.13)	(0.062)	(0.078)	(0.18)	(0.15)
Mean of dependent variable	0.326	0.001	0.332	0.076	0.105	0.576	0.319
Observations	267	267	267	267	267	267	267
	Panel F: 2019 and after – Reduced form						
Pleasant weather	0.025**	0.00025*	-0.021	-0.015**	0.0089*	-0.0052	-0.0037
	(0.011)	(0.00014)	(0.013)	(0.0058)	(0.0053)	(0.014)	(0.014)
Mean of dependent variable	0.326	0.001	0.332	0.076	0.105	0.576	0.319
Observations	267	267	267	267	267	267	267
Election fixed effects	Y	Y	Y	Y	Y	Y	Y
City-level controls	Y	Y	Y	Y	Y	Y	Y

Note: OLS and reduced form estimate of the effect of demonstrations on aggregated electoral outcomes, pooled over elections. Column 1-4 aggregate all parties along a left-right axis. Respectively, column 1 includes all left-wing parties, column 2 centrist parties, column 3 right-wing parties (including far-right) and column 4 far-right parties. Outcomes of columns 5 and 6 aggregate parties along an anti-corruption / not anti-corruption axis. Column 7 shows results for abstention. Panels A and B pool together all elections starting from 2011 (2011, 2015, 2016, April 2019 and November 2019 general elections, and 2014 and May 2019 European elections). Panels C and D pool only the elections starting in 2014, and panels E and F only the 2019 elections. Panels A, C and E show OLS estimates, panels B, D, F show reduced form estimates. All models include the usual city-level controls and election fixed effects. Standard errors (in parentheses) are clustered at the municipality level (*10%, **5%, ***1%).

Table E4: Vote for far-right in 2019 among voters of the right in 2016

	Vote far-right in April 2019 among voters for the right in 2016
	(1)
Ratio demonstrators	-0.842*
	(0.487)
Observations	1,122
City-level controls	Y
Individual controls	Y

Note: Effect of the ratio of demonstrators in the capital of a province, instrumented by unpleasant weather during the June 19, 2011 demonstration, on vote for the far-right in April 2019 among voters for the right in 2016. The sample comes from the CIS post-electoral survey for the April 2019, and is limited to respondents declaring that they voted for the right in the 2016 congressional elections. The model includes the usual city-level and individual-level controls. Standard errors (in parentheses) are clustered at the province level. (*10%, **5%, ***1%)

Table E5: First stage regression for the survey results.

	Ratio demonstrators		
	2011 and after	2014 and after	2019 and after
	(1)	(2)	(3)
Unpleasant weather	-0.0357***	-0.0353***	-0.0331***
	(0.00593)	(0.00597)	(0.00623)
Probability of unpleasant weather	0.00878	0.00718	-0.000123
	(0.0256)	(0.0254)	(0.0246)
City-level controls	Y	Y	Y
Individual controls	Y	Y	Y
Month fixed effects	Y	Y	Y
Mean of dependent variable	0.0589	0.0581	0.0537
F statistic for unpleasant weather	36.17	35.01	28.28
R^2	0.85	0.84	0.81
Observations	288,774	219,074	81,089

Note: First stage regression of the ratio of demonstrators in the June 19 protest to the city population on unpleasant weather during this same demonstration. The sample consists of all individuals interviewed in the monthly "CIS Barometer" between June 2011 and December 2019. Column 1 shows the result for the whole sample, column 2 for the sample of survey waves carried out in 2014 and later, and column 3 for survey waves in 2019. All models include the usual city-level controls and individual controls. Survey-wave fixed effects are also included (a new wave is conducted every month). Standard errors (in parentheses) are clustered at the province level (*10%, **5%, ***1%).

Table E6: Reduced form and OLS: Effect of pleasant weather on political opinions

OUTCOMES	Corruption main worry (1)	Political self-classification Left (2)	Right (3)	Far-right (4)	Don't know (5)
	Panel A: 2011 and after – OLS				
Ratio demonstrators	0.263***	0.892**	-0.450**	-0.133**	-0.442
	(0.0814)	(0.369)	(0.219)	(0.0651)	(0.361)
Mean of dependent variable	0.121	0.641	0.243	0.0299	0.116
Observations	285,356	258,525	258,525	258,525	258,525
	Panel B: 2011 and after – Reduced form				
Pleasant weather	0.0149***	0.0560**	-0.00185	-0.00971***	-0.0542*
	(0.00464)	(0.0254)	(0.00990)	(0.00361)	(0.0271)
Mean of dependent variable	0.121	0.641	0.243	0.0299	0.116
Observations	285,356	258,525	258,525	258,525	258,525
	Panel C: 2014 and after – OLS				
Ratio demonstrators	0.311***	0.963**	-0.583**	-0.166**	-0.380
	(0.0935)	(0.366)	(0.222)	(0.0647)	(0.345)
Mean of dependent variable	0.138	0.644	0.244	0.0309	0.112
Observations	218,314	200,727	200,727	200,727	200,727
	Panel D: 2014 and after – Reduced form				
Pleasant weather	0.0177***	0.0570**	-0.0121	-0.0103***	-0.0449
	(0.00528)	(0.0253)	(0.00998)	(0.00369)	(0.0270)
Mean of dependent variable	0.138	0.644	0.244	0.0309	0.112
Observations	218,314	200,727	200,727	200,727	200,727
	Panel E: 2019 and after – OLS				
Ratio demonstrators	0.242***	0.802**	-0.724***	-0.246***	-0.0781
	(0.0857)	(0.320)	(0.234)	(0.0790)	(0.295)
Mean of dependent variable	0.0892	0.646	0.259	0.0369	0.0944
Observations	80,649	75,898	75,898	75,898	75,898
	Panel F: 2019 and after – Reduced form				
Pleasant weather	0.0168**	0.0490**	-0.0131	-0.0114***	-0.0359
	(0.00689)	(0.0200)	(0.0102)	(0.00383)	(0.0230)
Mean of dependent variable	0.0892	0.646	0.259	0.0369	0.0944
Observations	80,649	75,898	75,898	75,898	75,898
Month fixed effects	Y	Y	Y	Y	Y
City-level controls	Y	Y	Y	Y	Y
Individual-level controls	Y	Y	Y	Y	Y

Note: OLS and reduced form estimates of the effect of the size the main demonstration in a province on concerns about corruption and on self-placement on the left-right axis (1-10). The sample consists of all individuals interviewed in the monthly "CIS Barometer" between June 2011 and December 2019. Column 1 shows results for responding that corruption is the main problem of Spain. Column 2-4 show the self-placement on the 1-10 left-right axis (left, right, far-right respectively). Column 5 shows the results for the "don't know" answer to the political self-placement question. Panels A and B pool together all interviews conducted from 2011. Panels C and D pool interviews from 2014 on, and panels E and F only interviews conducted during 2019. Panel A, C, E present OLS estimates. Panel B, D, F present reduced-form estimates of the effect of pleasant weather. All models include the usual city-level controls and individual controls. Survey-wave fixed effects are also included (a new wave is conducted every month). Standard errors (in parentheses) are clustered at the province level (*10%, **5%, ***1%).

Table E7: Percentage of 15M participators who started using (and had never used before) different online networks and platforms for 15M related communications.

Started using for 15M...	%	count
Twitter	32.6%	313
N-1	20.1%	193
Bambuser	19.8%	190
Pads (e.g. Google docs)	16.7%	160
Mumble	14.0%	134
Livestream	11.3%	108
Facebook	9.8%	94
Mailing lists	8.3%	80
Whatsapp	5.6%	54
Blogs	2.8%	27
Total		959

Note: Only the top 10 platform are shown. Data come from Redes, Movimientos y Tecnopolítica (2014), a survey of 15M participants.

Table E8: Tweets mentioning 15M among tweets mentioning Podemos

OUTCOMES	Ratio tweets 15M among tweets Podemos (1)
	Before 2015 election
Ratio demonstrators	0.75*
	(0.045)
Mean of dependent variable	0.001
City-level controls	Y
Observations	89

Note: Regression of the share of tweets about Podemos in the month before the 15M congressional elections that also mention 15M on the the ratio of the number of demonstrators over the population of the municipality. The model includes the usual city-level controls. Robust standard errors in parentheses (*10%, **5%, ***1%).

Table E9: Internal migration driven by protest differences

OUTCOMES	Migration from origin to destination provinces				
	Count (1)	Count over population (2)	Count over total emigrants (3)	Count over origin population and destination population (4)	Count over origin emigrants and destination immigrants (5)
Different weather	226.2	0.000556	0.00394	0.0192	0.0926
	(447.0)	(0.000395)	(0.00269)	(0.0174)	(0.130)
Distance	-14.68***	-1.96e-05***	-0.000160***	-0.00109***	-0.00893***
	(2.176)	(3.61e-06)	(2.01e-05)	(0.000112)	(0.000821)
Distance (squared)	0.00471***	6.96e-09***	5.32e-08***	3.99e-07***	3.07e-06***
	(0.000823)	(1.74e-09)	(9.83e-09)	(6.40e-08)	(4.71e-07)
Origin province fixed effects	Y	Y	Y	Y	Y
Destination province fixed effects	Y	Y	Y	Y	Y
Observations	2450	2450	2450	2450	2450

Note: Regression of measures of internal migration between pairs of provinces on whether the provinces had different weather during the demonstration of June 29th, 2011. The sample consists of all pairs of an origin province and a (different) destination province. Numbers of internal migrants aggregated over the second half of 2011 and 2012 to 2021. All models include origin province fixed effects and destination province fixed effects, and controls for the distance and the square of the distance between the capitals of the provinces. Column 1 uses as outcome the number of migrants from the origin to the destination province, column 2 the ratio of the number of migrants from the origin to the destination over the population of the origin province, column 3 the ratio of the number of migrants from the origin to the destination over total migrants from the origin, column 4 the ratio of the number of migrants from the origin to the destination over the population of the origin province and divided by the share of the total population of Spain represented by the destination province, and column 5 the ratio of the number of migrants over total emigrants from the origin province, divided by the ratio of immigrants in the destination province over the total number of internal migrants in Spain. Standard errors (in parentheses) are two-way clustered at the level of the origin province and the destination province (*10%, **5%, ***1%).

Table E10: Voter inertia for Podemos

OUTCOMES	Stopped voting Podemos among Podemos voters (1)	Started voting Podemos among Podemos non-voters (2)
	Panel A: Vote in 2016 depending on vote in 2015	
Pleasant weather	-1.53	0.07
	(3.44)	(0.75)
Mean of dep. var.	12.00 %	6.12 %
Observations	933 (17.33 %)	3630 (67.41 %)
	Panel B: Vote in April 2019 depending on vote in 2016	
Pleasant weather	2.75	1.64***
	(4.31)	(0.45)
Mean of dep. var.	30.96 %	3.00 %
Observations	743 (14.59 %)	4237 (83.23 %)
	Panel C: Vote in November 2019 depending on vote in April 2019	
Pleasant weather	3.95	1.88***
	(4.00)	(0.54)
Mean of dep. var.	17.04 %	2.35 %
Observations	493 (11.96 %)	3064 (74.33 %)
City-level controls	Y	Y
Individual controls	Y	Y

Note: Regression of a dummy variable indicating whether an individual switched to voting for Podemos or switched to voting for an other option on pleasant weather in the capital of the individual's province. The sample comes from the CIS post-electoral surveys. Panel A uses the 2016 post-electoral survey, panel B the April 2019 survey and panel C the November 2019 survey. Column (1) restricts the sample to voters having voted Podemos in the previous election, and the outcome is 1 if the individual votes for another party in the election of interest. Column (2) restricts the sample to individual having voted for another party and the outcome is 1 if the individual votes for Podemos. Coefficients have been multiplied by 100 for presentation and can be interpreted as percentage points. The model includes the usual city-level and individual-level controls. Standard errors (in parentheses) are clustered at the province level (*10%, **5%, ***1%).