A Appendix

A.1 Replication data

All data required for the replication of quantative analyses will be made available on the Harvard Dataverse. I will also provide these data to reviewers upon request.

A.2 Twitter data

Twitter data comes from the 1% Twitter streaming API for the hashtag `#sidibouzid`. This captures 1% of all tweets (from any location) using this hashtag. Data were collected during the course of the uprising by the editor of Nawaat Sami Ben Gharbia and provided to the author in person during a fieldwork stay in Tunis in April, 2017. These data comprise some 85 thousand tweets over the period 18/12/2010-14/11/2011. Data collection only began after the first day of protests in Sidi Bouzid, thus explaining the lack of data for December 17. For the purposes of the text analysis in Section 3 of this article, I employed the R package *arabicStemR* authored by Rich Nielsen, which allows users to stem, clean, and transliterate Arabic text(see `https://cran.r-project.org/web/packages/arabicStemR/arabicStemR.pdf`). By transliterating Arabic text into Latin script, the user is able to conduct textual analysis within one sample. Words used in these analyses are listed in Table A.1. Words were selected by identifying terms that shared the English, French, or Arabic roots for "election" or "democracy" words; i.e., "democ*," "elect*," "démoc*," "élect*," "dymqrat*," "antkab*." Day 29 (January 14) was removed as by this point Ben Ali's ouster was secured, prompting many to begin speaking about democracy and elections, thereby spuriously inflating democracy talk. Our interest here is, instead, the contents of online activity before any outcomes were known.

Table A.1: Democracy words in `#sidibouzid` data. Note: arabicStemR converts quotation marks to "quot" thus explaining the appearance of this suffix on certain of the words.

Words	alantkabat	balantkabat
democracy	d'elections	bantkabat
elections	wantkabat	d'élection
antkabat	élection	democracies
democratic	alantkaby0	democracymyass
aldymqraty	aldymqratyyn	democrasy
aldymqraty0	baldymqraty0	democrates
lantkabat	democracynow	democratia
llantkabat	democrat	democratically
élections	democratique	democratization
	dymqraty0	democratize
	wdymqraty0	democrazia
	89lections_en_tunisie	democráticas
	antkab	electionshttp
	atheniandemocra	electionslibresentunisie
		l'election
		lldymqraty
		sidibouzidalantkabat

A.2.1 Twitter location data

One concern about the Twitter data used in this article is that it will contain tweets that do not necessarily originate from Tunisia. This means that any inference we make from over-time trends in word frequency is necessarily limited and may not reflect the type of information and grievance circulating within online Tunisian communities. I check this in the following way: I first requested the user profile information for accounts present in the Tunisian Revolution Twitter data (who are still active on Twitter) through the Academic Research Product Track API. I then take those accounts that do have some location information and I pass these through the OpenStreetMaps API using the `tidygeocoder` R package (Cambon et al., 2021). I am then able to determine whether the user likely resides in Tunisia. Figure A.1 plots the locations and relative intensity of user activity (number of tweets).
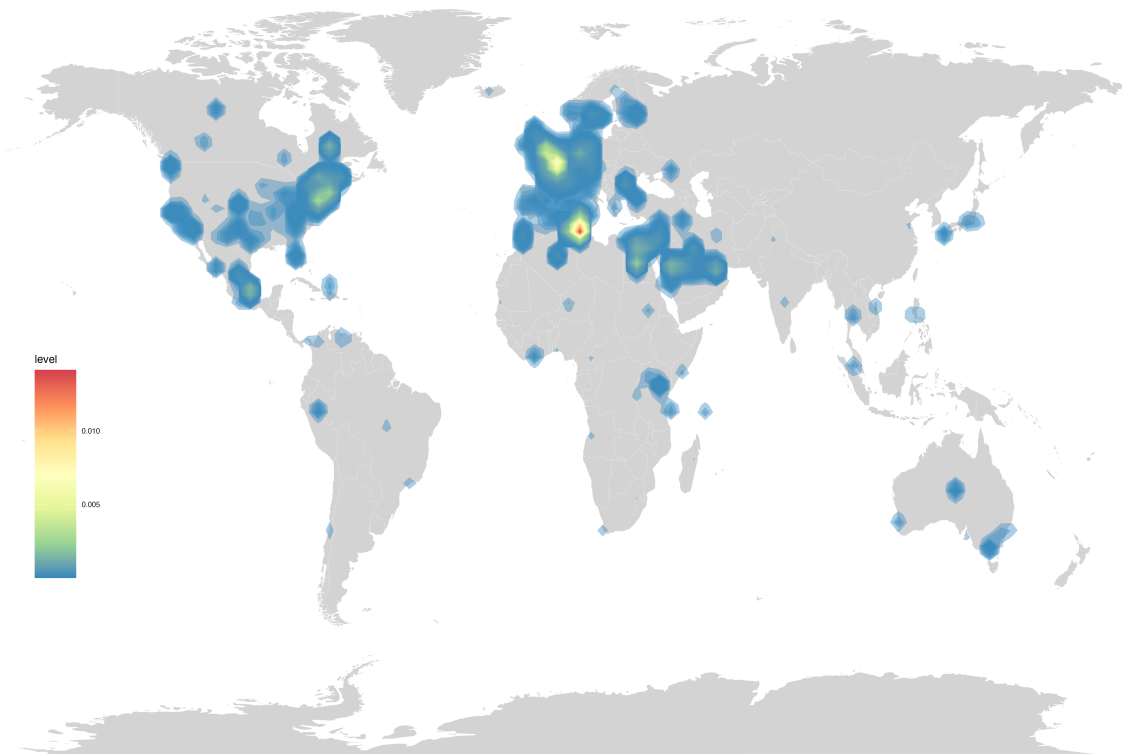
Figure A.1: Map of locations of tweeters inferred using `tidygeocoder` and relative frequency represented by 2D kernel density estimation procedure.

Having inferred user locations, I can filter the original Twitter data to remove any tweet by an author that isn't among this subset of users. I then redo the Twitter word frequency analyses using only these data. The trends are less obvious and definitive given the reduced size of the data but point in the same direction.
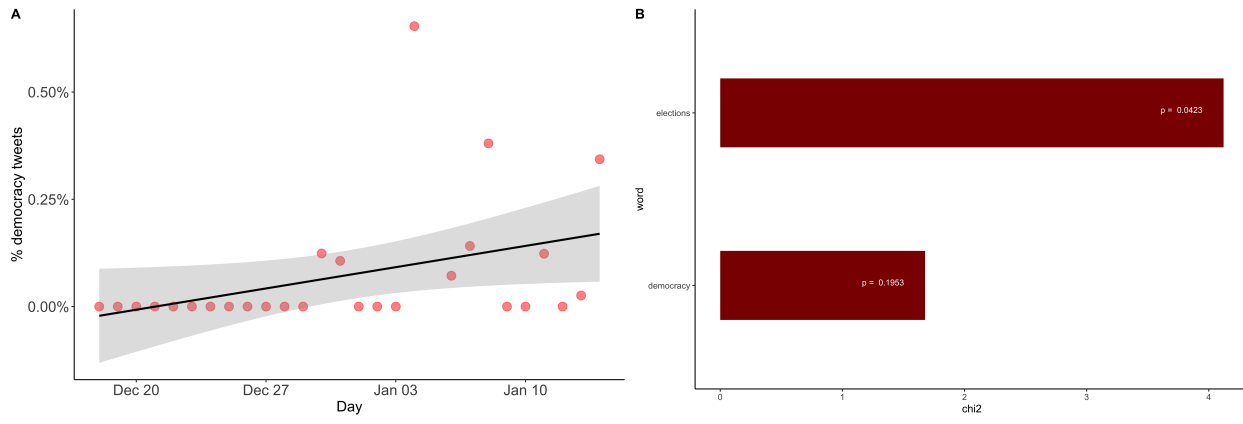

Figure A.2: **A** Frequency of words related to democracy in `#sidibouzid` Tunisia data (% of total); **B** Relative democracy word keyness by Stage (1 versus 3)

A.3 News data

The news articles come from a subset of news media data obtained by the author from the founder-creator of news aggregation platform `turess.com`, Mossaab Bagdouri. These data come in the form of the raw text of news articles from thirteen different news sources: "akhbar", "alchourouk", "alfajrnews", "alhiwar", "almasdar", "alwasat", "assabah", "assarih", "attounissia", "babnet", "echaab", "kalima", and "tap." I filter these news articles to take only those articles classified (on the `turess.com` platform) as domestic politics articles. After filtering in this way, there are 2375 separate news articles over the period December 17–January 14. The news data were pre-processed to remove non-Arabic words, numbers, and punctuation. The full corpus from which they deriver are described in full at: https://osf.io/68zn4/.

A.4 Nightlights data

Nighttime lights data for the Visible Infrared Imaging Radiometer Suite Day/Night Band Nighttime Lights (VIIRS DNB) are retrieved from `https://ngdc.noaa.gov/eog/viirs/download_dnb_composites.html` and for the Defense Meteorological Satellite Program Operational Linescan System Global Radiance Calibrated Nighttime Lights (DMSP-OLS GRCN) from `https://ngdc.noaa.gov/eog/dmsp/download_radcal.html`, last accessed: October 3, 2019. VIIRS data should be preferred over DMSP-OLS data due to known problems of pixel saturation and low resolution in the latter (Elvidge et al. 2013). In order to calculate delegation-level values of light intensity, the QGIS mapping software was used. A shapefile of each of Tunisia's 264 delegations was clipped to the GEOTIFF files in which the VIIRS data are supplied. Using the "zonal statistics" plugin, mean pixel values for each geographical unit (delegation) were then calculated. Mean intensity is calculated as the sum of pixel values denominated by the total number of pixels in a given delegation. Unlike the DMSP-OLS satellite, which are measured on a 0-63 relative scale, VIIRS DNB pixel radiance is calculated as $\frac{nanoWatts}{(cm^2 sr)}$, and does not suffer from known problems of pixel saturation as a result of scale limitation (Elvidge et al. 2013; Wu and Wang 2019). The VIIRS data are only available from April 2012 onwards. This evidently postdates the revolution. In order to compare against a light intensity development proxy recorded from before the revolution, I rely on GRCN data. The correlation between the log-transformed values of mean pixel values and IDR by delegation for each of these satellite data are displayed below in Figure A.3 below. The strength of the correlation with IDR for each measure is almost identical at $r{\sim}.8$.
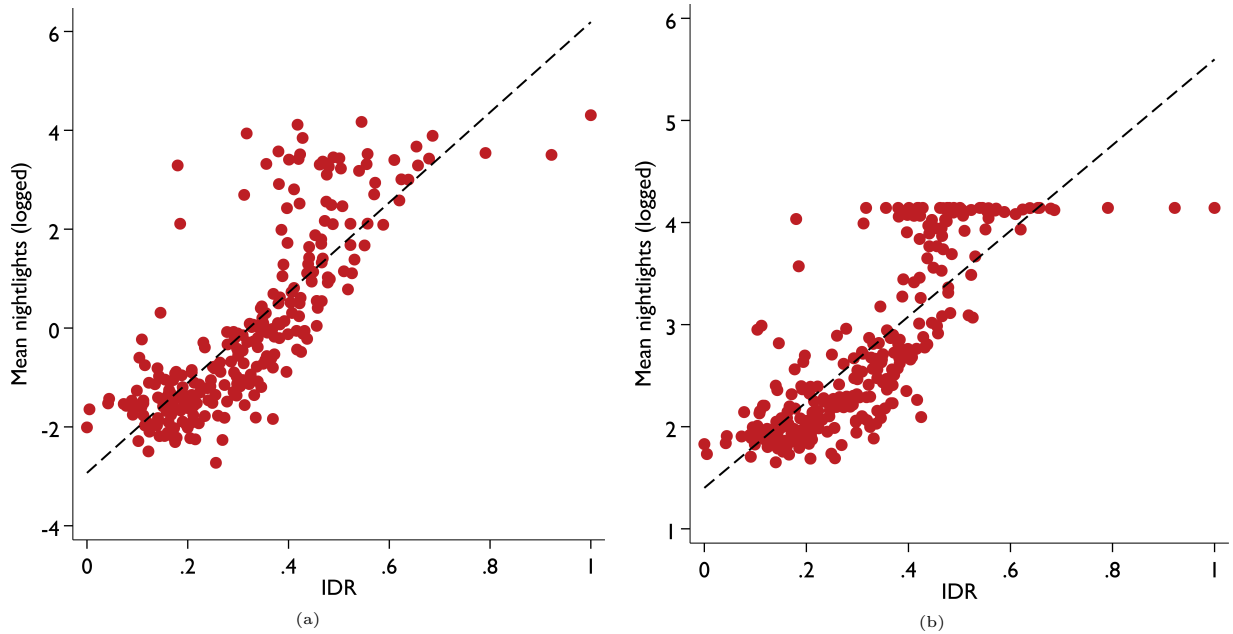
Figure A.3: (a) Correlation between delegation-level VIIRS DNB satellite data (logged) and IDR; (b) Correlation between delegation-level DMSP-OLS GRCN data (logged) and IDR)

A.5 Event data sources

While the majority of protest event analyses derive their data from newspaper reports, this was not possible during for the twenty-nine day revolutionary uprising in Tunisia where a media blackout aimed specifically at stymieing the flow of information. I overcome this obstacle by triangulating multiple alternative sources of information. I made use first of Facebook pages, two of which were already in operation before the uprising, and two of which were created for the specific purpose of posting news of protests. The names and details of these pages are below. The pages were archived in PDF format, retaining the link structure, thereby enabling the manual coding of protest events, photos, and videos from each of them. These sources were then supplemented with multiple further sources of information, including national newspapers, international newspapers, a post-revolutionary investigatory commission, an archive of protest reports from the dissident blog *Nawaat*, as well as the *Archive Numérique de la Révolution*—a digital archive at the Tunisian National Archives devoted to the collection and archiving of photos and videos from the period of the uprising. The groups used, as well as further sources used for constructing the event catalogue for the revolutionary period, are listed below. Table A.2 details the number of protest reports that derive from each source for the full event catalogue (EC) dataset and the event history (EH) dataset, which only records information for first protest occurrence in a given delegation. Only the EH dataset is used in the analyses for this article.

A.5.1 Facebook

- شعب تونس يحرق في روحو يا سيادة الرئيس *Chaab Tunis Yahriq fi Ruhu ya Siyyadat al-Ra'is* (Mr. President, the People of Tunisia are On Fire (PTON)). This page was set up upon the outbreak of protests, as the title of the group suggests. It went through six iterations as it was continually hijacked by unknown cyber attackers, likely linked to the Ben Ali regime. When hijacked, the founders would set up a new page by the same title but with the number of the version of the group appended. The pages contained information on protest for everyday of the uprising with the exception of the period 02/01/11-08/01/11 when the page was down. Protest reports would often report on the type of protest (e.g., march, occupation, demonstration), include some mention of size (e.g. 'a group of', 'large', 'huge'), and give some mention of source (most often 'union sources'). When the report cited 'unconfirmed reports', the report was not included in the event catalogue. In total these pages provided information on 193 protest events, 140 of which were corroborated with at least one other source. Out of the 193 protest events recorded from this source, 94 could be checked against video evidence. Figure A.4a gives an example of a video protest report posted to this page.

Table A.2: Number of reports by source in event catalogue, full event catalogue (EC) and event history (EH) dataset.

Source	Freq. EC	Freq. EH	Source	Freq. EC	Freq. EH
AFP	8	1	Kalima Tunis	343	87
Al-Jazeera News	8	3	Le Monde	1	-
Al-Jazeera Video	18	3	Le Point	2	-
Al-Badil	1	-	MAMN (FB)	33	6
Al-Wasat	1	1	NATS (FB)	195	48
ANR	63	10	Nawaat	144	35
ANR Video	61	10	Nawaat Video	106	27
BBC	1	-	Paris Match	1	-
Bouderbala	29	7	Photo	25	4
BusinessNews	4	2	PTON (FB)	193	68
Chourouk	85	13	Reuters Arabic	4	1
Elaph	1	-	Takriz (FB)	42	22
France24	2	1	Video	106	41
France24 Video	2	1			

(a)

(b)

Figure A.4: (a) Typical video protest report from PTON. Text underneath reads: 31/12/10: Er-Rouhia—Siliana Governorate. Subsequent lines detail chants heard during the protest, including: "Don't be a coward, go out onto the street!"; "Work, freedom, national dignity!"; "Work is a right, you gang of thieves!", "No to tyranny and corrupt government!".; (b) Typical protest report from NATS. Text reads: "Kasserine, today 06/01/2011: A large march of students, teachers, unemployed and unemployed graduates left this morning from the UGTT offices, circling the city before stopping outside the RCD office at which point the chant "Down with the Dostour Party, down with the executioner of people" was heard alongside other chants calling for work, freedom, and national dignity. Lawyers also joined in the protest as the march went past the court building, at which point the police forcefully intervened. We will keep you updated".

- وكالة أنباء تحرّكات الشارع التونسي *Wikalat Anba' Taharrukat al-Shari'a al-Tunisi* (News Agency of the Tunisian Street (NATS)). This page was also set up upon the outbreak of protests in Sidi Bouzid and protest reports followed a similar formula to the above page. It has two iterations, after the first was also hijacked. The pages contained information on protest for the period 02/01/11-14/01/11. In total, this page provided information on 195 protest events, of which 106 could be matched with a secondary source. Out of the 195 protest events recorded from this source, 71 could be checked against video evidence. Figure A.4b gives an example of a typical protest report from this page.

- ما عندي مانقلك *Ma 'Ayndi Mankolek* (I Have Nothing to Say to You). This page was already in existence prior to the revolution. It was a forum for dissident opinion and took positions against, for example, online censorship. This page was active for the entire period of the uprising. In total, this page provided information on 33 protest events, 29 of which could be matched with a secondary source, and 18 of which could be checked against video evidence.

- **TAKRIZ** (Ball-Breaker). This page was already in existence prior to the revolution and ran alongside the now-defunct website *takriz.com*, founded in 1998. It was a forum for dissident opinion, anti-censorship activism, and often irreverent commentary on Tunisian affairs. Upon the outbreak of the revolution, its administrators began posting videos and reports of protests, encouraging their members to go out and protest at the same time. Unfortunately, the page was not available for the period 02/1/11-14/01/11. In total, this provided information on 42 protest events, of which 30 could be verified against a secondary source.

A.5.2 Radio/National News

- راديو كلمة تونس *Radio Kalima Tunis* (Word of Tunisia Radio). Kalima Tunis is a radio station set up in 2008 by journalist and human rights campaigner Sihem Bensedrine, and Omar Mestiri. The radio station operated from France after being banned in 2009, but still had journalists on the ground. It published multiple dispatches daily on the unfolding of protest events that have been archived by online news aggregator `turess.com`. In total, Kalima Tunis provided information on 342 protest events, of which 199 could corroborated with a secondary source.

- الشروق *Al Chourouk* (Sunrise). Pro-regime newspaper that only began to report on protest in the closing stages of the uprising. Nonetheless, it provided information on 86 protest events, of which 44 could be checked against a secondary source. Copies of these reports were obtained from `turess.com`. These articles were also checked against archived paper copies of the newspaper in the Centre de Documentation Nationale archive in Tunisia in order to check for any omissions on the `turess.com` website. No significant omissions were found.

A.5.3 Nawaat Archive

- Over the course of the revolution, *Nawaat*, a dissident blog established in 2004, reported on protest events. Most of these reports became unavailable after the shutdown of blogging host platform `posterous.com`. Using the Wayback Machine Internet Archive, however, I was able to retrieve this content. These yielded reports on 140 protest events, 127 of which could be triangulated with other sources.

A.5.4 Digital Archive of the Revolution

- A cache of videos and photos, collected by Jean-Marc Salmon in conjunction with archivists at the Tunisian National Archives, was handed over to the National Archives in Tunis in March 2017. Reporting on the collection can be accessed here: `https://preview.tinyurl.com/y2odrjpn`. The collection contains around 1000 photos and 800 videos. Researchers focused their efforts on Greater Tunis, Kasserine, and Sidi Bouzid meaning that documents from other regions may have yet to find their way into the archive. Nonetheless, the archive contains documents for each of Tunisia's twenty-four governorates, and nearly all photos and videos are datestamped. I searched through each date for every district in every governorate over the period of the revolution and triangulated protest reports with the event catalogue, updating e.g., protest sizes according to photo and video availability, and adding new events when omitted. The archive yielded reports on 63 protest events that could reliably be located in a delegation, and 40 of these could be triangulated with other sources.

A.5.5   Videos and Photos

- Videos. Videos were often posted to the Facebook groups listed above accompanied with reports on the protest itself. In total 222 videos were found, all of which were matched with a protest report. Videos were nearly all accompanied by a date in the comments under the video. Alternatively, videos would be accompanied by a comment such as "Situation today in [name of town/city]", meaning the date could then be assumed as the date of the posting itself.

- Photos. Photos were also often posted to the Facebook groups listed above accompanied with reports on the protest itself. In total 25 photos were used and matched with protest reports. Photos would also either be accompanied by a date or the date of the event could be deduced from comments in the post e.g. "Photos from protest yesterday in [name of town/city]".

A.5.6   International News

- International news sources were also used when available. Two of these took the form of evening news reports, archived versions of which were posted on the Facebook pages described above. The first of these is *Al Jazeera: The Maghreb Harvest* (الجزيرة: الحصاد المغربي). This was a daily news round up of events in Arab North Africa that, given the ongoing protests in Tunisia, focused primarily on Tunisian affairs over the twenty-nine days of the uprising. *France 24* (24 فرانس) also provided a daily round of news in North Africa, with a particular focus on Tunisia over the twenty-nine days of the revolt. News articles from Arabic-, French-, and English-language news outlets were also archived when posted on the Facebook groups listed above. These included: *BBC News*, *Agence France Presse*, *Le Point*, *Le Monde*, *Reuters Arabic*, *Agence Tunis Afrique Presse*, *Business News Tunisia*. Further articles from international news media were extracted from Google News. These were then all matched against existing protest reports. In total, these sources together provided information on 53 separate protest events, all of which were verified against a secondary source.

A.5.7   Bouderbala Commission

- The full report is available online here: `http://www.leaders.com.tn/uploads/FCK_files/Rapport%20Bouderbala.pdf` In this first report, the Commission looked into abuses committed from 17 December 2010 up to the first elections on 23 October 2011. Bouderbala is himself a lawyer and Honourary President of the Tunisian Human Rights League (Ligue tunisienne des droits de l'homme). The list of deaths published was described as "provisional" and the Commission has yet to publish its final version. While the existing list may only be provisional, and it does not identify police responsible, the report is nonetheless comprehensive (running to over 1,000 pages) and provides a rich source of information for the purposes of this study. The Bouderbala Commission verified deaths with visits to the homes of the bereaved and checked reports against available medical records, thus providing a confident estimate of levels of repression witnessed during protests. The reports included the circumstances of the death, the date of the incident, and the institutional identity of the perpetrator. Only those that stated explicitly that the individual was killed at the hands of state security (police, national guard or army), or was killed during a protest, were included for the analysis.

A.6   Event catalogue codebook

The variables, and their codings, for the event catalogue are listed below. Note that not all variables are used in the abaove analysis. They are listed here for completeness. Relevant variables for the present article include: *date*; *governorate*; *delegation*; *noofkilled*; *source*.

1. Date (*date*): The date of the protest occurring. This could be deduced from the report itself and the date attached to the post in Facebook, for example, or the date of the article posted on the website of the Radio Station or newspaper.

2. Governorate (*governorate*): The Governorate (one of 24) in which the protest took place. These would normally be reported in the protest reports themselves. Otherwise, reports would include reference to the specific town or area in which the protest took place, from which one could deduce the Governorate from the list of delegations in each Governorate or by Google searching the name of the town/city.

3. Delegation (*delegation*): The Delegation (one of 264) in which the protest took place. Normally the delegation could be deduced by Google searching the specific location of the protest (e.g. name of school, *Wilaya* building, UGTT building) or would be specifically reported.

4. Repertoire (*repert*): The type of protest. This was normally contained within the protest report itself or could be identified in the videos or photos of the protests. Repertoire could be one of "demonstration" (manifestation/protestation or مظاهرة/احتجاج); "march" (marche or مسيرة); "strike" strike (grève or إضراب); "general strike" (grève régionale or إضراب عام محلي); "blockade" (bloquer la route or قطع الطريق); attack (attaque or مهاجمة); "occupation" (occupation or اعتصام) or "sit-in" (sit-in or وقفة احتجاجية).

5. Protest location (*protlocation*): The specific location of the protest e.g. outside Wilayat building or UGTT offices.

6. Starting location (*protstartlocation*): A general identifier for the start location. One of "city centre"; "govt building" (e.g. any official building such as police station, local government building, RCD (ruling party) offices); "hospital"; "main road"; "public transport" (e.g. railway/bus station); "residential street"; "saha" (square); "school" or "university".

7. Moving to (*movingto*): Where the protest moved to, if it did move.

8. End location (*protendlocation*): General identifier for where the protest ended, from among those listed in variable #8.

9. Organizer (*organizer*): Organizer of the protest. This was coded for both those specifically identified as organizing the protest e.g. "unionists" normally from the national trade union federation the UGTT, or it was coded for the principal participants e.g. "teachers" or "students".

10. Number killed (*noofkilled*): The number killed during the protest according to the data provided by the Bouderbala Commission.

11. Source (*source*): Source of the protest report (e.g., names of Facebook group, newspaper, radio station etc.).

A.7   Interview data

I conducted a set of over forty interviews in the spring of 2017 and 2019 during two fieldwork visits in Tunisia. Interviews were with a diverse range of participants, including former members of student protest movements, trade unionists, unaffiliated (former) school students, and members of the one semi-legal opposition party—the Parti démocrate progressiste (PDP).

Interviewees were enrolled using a snowball sampling procedure, exploiting multiple entry points, and took place in Sfax, Sidi Bouzid, and Tunis—three regions that saw significant protest at different stages of the Revolution. The choice of field sites was informed by the event data: I wanted to visit sites that saw protest at distinct stages of the uprising. Some regions, such as Kasserine and Thala fell within this sampling framework but were not accessible at the time of the fieldwork visits due to security concerns. Interviews were variously conducted in Arabic, French or English and were semi-structured in format. Most interviews were recorded (and subsequently transcribed in full) unless informants desired otherwise. Names have been changed to preserve anonymity. Translations are all my own.

A.8   Methodological appendix—Ecological data sources

- For the development index used in the main event history analyses, I assemble data from the Comparative Study on Tunisian Regional Development (*Étude comparative en terme de développement régional de la Tunisie*) conducted by the Tunisian Competition and Quantitative Studies Institute (*Institut Tunisien de la Compétitivité et des Études Quantitatives*), a subsidiary of the Ministry of Planning and Regional Development (*Ministère du Développement Régional et de la Planification*). An example overview of the index and its measures (for the Beja governorate) can be accessed here: http://www.mpci.gov.tn/tn/Gov/indica/beja.pdf. The study was conducted in order to provide a composite delegation-level index of regional development (IDR). This in view of the considerable regional disparities that exist in Tunisia, particularly between the interior and the Sahel. The index ranges from 0-1, with 1 representing the most developed and 0 the least developed. The index was constructed using principal components analysis on twenty-six separate variables. A full list of measures used for the construction of the index, as well as links to further information, where available, is provided in Table A.3.

Table A.3: Sources and measurements of IDR data

Variable	Source	Measurement	Further information
Proportion roads classified (2010)	Direction générale d'infrastructure	proportion of roads (in Kms) classified over total Kms of road in delegation.	
Distance to city hubs	Direction générale d'infrastructure	distance between delegation and each of Tunisia's main three city hubs (Tunis, Sousse, Sfax).	
Distance to nearest port	Direction générale d'infrastructure	distance between delegation and nearest port/airport.	
Agricultural labour force (2004)	2004 Census	percentage individuals employed in agriculture denominated by total active population aged 15 or over.	http://nada.ins.tn/index.php/catalog/8
Sewage connection (2004)	2004 Census	percentage households connected to sewage system.	http://nada.ins.tn/index.php/catalog/8
Drinking water (2004)	2004 Census	percentage households with clean drinking water.	http://nada.ins.tn/index.php/catalog/8
Number doctors (2010)	Commissariat et offices, Tunisie	number of doctors per 1000 residents.	
Number pharmacies (2010)	Commissariat et offices, Tunisie	number of pharmacies per 1000 residents.	
Number beds (2010)	Commissariat et offices, Tunisie	number of hospital beds per 1000 residents.	
Number families "in need" (2011)	Commissariat et offices, Tunisie	number of individuals per 1000 residents in receipt of government assistance.	http://www.cres.tn
Poverty rate (2005)	Institut National de la Statistique, Tunisie	percentage households living below poverty rate.	http://www.ins.tn/fr/publication/enqu%C3%AAte-sur-la-consommation-des-m%C3%A9nages-2005-volume
Dependency rate (2004)	Institut National de la Statistique, Tunisie	percentage population 0-14 and 65+ denominated by total population.	http://nada.ins.tn/index.php/catalog/8
Population (2010)	Institut National de la Statistique, Tunisie	number of individuals in delegation.	
Population growth (2008-2012)	Institut National de la Statistique, Tunisie	average growth in population over the period 2008-2012.	
Net migration (2004)	Institut National de la Statistique, Tunisie	net migration over period 1999-2004.	http://nada.ins.tn/index.php/catalog/8
Educated population (2004)	Institut National de la Statistique, Tunisie	percentage population with secondary or higher education certificate denominated by total population.	http://nada.ins.tn/index.php/catalog/8
Illiteracy rate (2004)	Institut National de la Statistique, Tunisie	percentage population illiterate denominated by total population.	http://nada.ins.tn/index.php/catalog/8
Bac admissions (2010)	Commissariat et offices, Tunisie	rate of admission to the Baccalaureate (final school examinations required for university admissions).	
Labour supply (2012)	Direction générale des ressources humaines	job opportunities per 1000 residents.	
Labour demand (2012)	Direction générale des ressources humaines	job demands per 1000 residents.	
Placement rate (2012)	Direction générale des ressources humaines		
Private industry (2010)	Institut National de la Statistique, Tunisie	number private enterprises per 1000 residents	http://www.ins.nat.tn/fr/themes/entreprises
Herfindahl index (2010)	Institut National de la Statistique, Tunisie	sum of squares of (delegation-level) market shares of enterprises	
Average private business size (2010)	Institut National de la Statistique, Tunisie		
Employed in private sector (2010)	Institut National de la Statistique, Tunisie	percentage employed in private sector	
Unemployment rate (2004)	Institut National de la Statistique, Tunisie	percentage individuals aged 18-59 without employment denominated by total active population aged 18-59.	http://nada.ins.tn/index.php/catalog/8

- The variable measuring youth population was taken from the 2014 census. The measurement of each of this item, as well as several others used in supplementary analyses detailed below, are listed in Table A.4.

- For a supplementary analysis, results of which below in Table A.7, data at the delegation level was taken from publicly available census data from censuses conducted in 2004 and 2014. From these, I extracted, from the 2004 census, data on delegation-level illiteracy rates. From the 2014 census, I extracted delegation-level graduate unemployment rates, illiteracy, youth population, % employed in agriculture, and % internet use. The measurement of each of these items, all taken at the delegation level, is listed in Table A.4.

Table A.4: Sources and measurements of ecological data

Variable	Source	Measurement	Further information
Illiteracy rate (2004)	2004 Census	percentage illiterate individuals aged 10 or over denominated by total population aged 10 or over.	`http://nada.ins.tn/` `index.php/catalog/8`
Youth population (2004)	2004 Census	percentage individuals of given age bracket (either 15-19 or 20-29) denominated by total population.	`http://nada.ins.tn/` `index.php/catalog/8`
Graduate unemployment rate (2014)	2014 Census	percentage graduates without employment denominated by total graduate population.	`http://census.ins.tn/` `fr/resultats`
Illiteracy rate (2014)	2014 Census	percentage illiterate individuals aged 10 or over denominated by total population aged 10 or over.	`http://census.ins.tn/` `fr/resultats`
Youth population (2014)	2014 Census	percentage individuals of given age bracket (either 15-19 or 20-29) denominated by total population.	`http://census.ins.tn/` `fr/resultats`
Internet use (2014)	2014 Census	percentage individuals aged 10 or over who use the internet denominated by total population aged 10 or over.	`http://census.ins.tn/` `fr/resultats`

- There are significant time intervals between the taking of the censuses and the outbreak of the Tunisian Revolution. In order to get a handle on this, I also extracted data from a 2010 Labour Survey in Tunisia (*Enquête nationale sur l'emploi 2010*), available here: `http://www.ins.tn/fr/publication/enqu%C3%AAte-nationale-sur-l%E2%80%99emploi-2010`. Figures from 2010 are only available at the governorate level. To adjudicate between measures, I am then able to compare correlations between the 2004 and 2014 governorate-level data and that available at the governorate level from 2010. These are displayed in Table A.5 below. As should be clear, most measures are highly correlated between the two measurement periods. Youth population is strikingly weak, however. It is also difficult to adjudicate between the 2004 and 2014 measures as both are in turn highly correlated, at the governorate level, with the measure taken in 2010. For each of the models in Table 1, ecological measures are taken from 2014, being more temporally proximate. Rerunning the main and supplementary analyses in Table A.7 below with ecological data taken from 2004 gives identical results. The youth population measure from 2004, however, loses significance. I justify the use of the 2014 measure as it is more temporally proximate than that for 2004.

- The diffusion variable used in the main event history analysis takes the following functional form:

$$P_{1it} = \sum_{\tau=t-1}^{t-1} \sum_{j=1}^{J} \frac{P_{j\tau}}{\sqrt{d_{ij}}}$$

, where $d_{ij}^{-0.5}$ represents the square root of distance between the centroid of delegation i and delegation j and $P_{j\tau}$ represents a binary indicator of whether delegation j witnessed any protest at time τ. These variables were generated

Table A.5: Comparisons of ecological data

Variable	Level of measurement	Years compared	Bivariate correlation
Illiteracy rate	Delegation	2004 & 2014	0.97
Youth population	Delegation	2004 & 2014	0.31
Graduate unemployment rate	Governorate	2010 & 2014	0.95
Internet use	Governorate	2010 & 2014	0.68
Youth population	Governorate	2004 & 2010	0.81
Youth population	Governorate	2010 & 2014	0.81

with Keisuke Kondo's *spgen* set of commands (2017). A distance of 15 km was used to capture nearby protest—the mean distance between two polygon centroids is 14.02 km. Varying this buffer from 10-40km does not substantively alter findings.

A.9 Methodological appendix—Survey data

The survey data used in the analysis is taken from Wave II of the Arab Barometer, a widely used set of nationally representative surveys from the MENA region. An extra battery of questions were added in 2011 for Tunisia (and Egypt) relating to participation in Arab Spring protests and their aftermath. The Arab Barometer is a joint initiative of the University of Michigan, Princeton University, and the Arab Reform Initiative. The funding for Wave II of the surveys came from the Canadian International Research and Development Centre (IDRC), the United Nations Development Program (UNDP), and the United States Institute of Peace (USIP). The fieldwork for the Tunisia survey was conducted from September 30, 2011-October 11, 2011. The Principal Researcher was Iman Mizlini of Sigma Conseil, an independent polling organization in Tunisia (see http://www.sigma.tn/Fr/accueil_46_9). The sample consisted of Tunisian adults aged 18 or over, stratified by governorate and urban/rural. Surveys were conducted face-to-face in Arabic. The survey used a three-stage stratification process, with delegations selected proportional to population size, sectors (within delegations) selected proportional to size, and blocks of sectors then selected at random. A block consisted of ten households. Individuals within households were then randomly selected using a Kish table. The sample is self-weighted but includes post-stratification weights to correct for age and gender imbalances.

The model by Doherty and Schraeder (2018) that I replicate includes measures for individual income (logged); education (six categories of schooling); age (years); gender (male=1); employment status (binary); retired/housewife (binary); student (binary); religiosity (index based on three questions relating to prayer, Qur'an reading, and self-reported religiosity); mosque attendance (five items of regularity of attendance); whether district of residence rural (binary); interest in politics (four items of level of interest); whether friends participated in protests (binary); whether member of any civil society group (binary). The model also included governorate indicators for all twenty-four governorates, and used post-stratification weights provided with the dataset.

A.9.1 Survey questions and codings:

- Protest participation (T902): "Did you participate in the protests against former president Zain Al-Abdeen Ben Ali between December 17th, 2010 and January 14th, 2011?" [Yes=1, No=0].

- Timing of protest participation (T903): "If you answered "yes" to question 902 (regarding your participation in the protests): Did you participate in any of the following protests?":

1. December 17, 2010 to January 1, 2011 [Yes=1, No=0].
2. January 2 –January 9, 2011 [Yes=1, No=0].
3. January 10 –January 14, 2011 [Yes=1, No=0].
   From these questions, the ordinal participation outcome measure was generated. This was coded 1 if the respondent participated from stage 1 (i.e., responded in the affirmative to item 1 of T903), 2 if the respondent protested from stage 2 (i.e., responded in the affirmative to item 2 of T903 and not to item 1), 3 if the respondent first protested at stage 3 (i.e., responded in the affirmative to item 3 of T903 and not to items 1 and 2), and 4 if the respondent participated at no stage (i.e., responded in the negative to T902). A second outcome measure, used for the multinomial logistic analysis, was coded identically except those who never protested were coded 0 and constituted the base reference outcome.

- Income (Q1014): Respondent was asked to report "Income including all wages, salaries and rent allowances". This was recorded in 2011 US dollars.

- Education (T1003): Respondent was asked to report highest level of education. Six options were offered: "1: Illiterate; 2: Elementary; 3: Preparatory/Basic; 4: Secondary; 5: BA; 6: MA and above". The variable used retained all items in an ascending six point scale.

- Age (Q1001): Respondent was asked to report age. Recorded in years.

- Male (Q1002): Gender of respondent recorded by interviewer. [Male=1, Female=0].

- Employed (Q1004): Respondent was asked "Do you work?" [Yes=1, No=0].

- Student/Out of workforce (Q1005): Respondent was asked "Are you: 1: Retired; 2: A housewife; 3: A student; 5: Unemployed (looking for work); 4: Other (specify)". Here, a separate binary student variable was coded, a variable for being out of thr workforce (1 if retired/housewife), and a binary variable for being unemployed. The latter (unemployed) is the reference category in the full model. Only three respondents gave other categories. These were coded as missing.

- Religiosity (Q609, T6101, T6106): Composite index constructed using responses to three questions. A first (Q609) asks respondents: "Generally speaking, would you describe yourself as...? 1: Religious; 2: Somewhat religious; 3: Not religious". A second (T6101) asks respondents: "Do you pray daily? 1: Always; 2: Most of the time; 3: Sometimes; 4: Rarely; 5: Never". A third (T6106) asks respondents: "Do you listen to or read the Quran/the Bible? 1: Always; 2: Most of the time; 3: Sometimes; 4: Rarely; 5: Never". Responses were rescaled so that higher numbers indicated greater religiosity. The item was then standardized with mean 0 and standard deviation 1.

- Mosque attendance (T6105): Respondent was asked: "Do you attend Friday prayer/Sunday services? 1: Always; 2: Most of the time; 3: Sometimes; 4: Rarely; 5: Never". Responses were rescaled so that higher values indicated greater frequency of attendance.

- Interest in politics (Q404): Respondent was asked: "In general, to what extent are you interested in politics? 1: Very interested; 2: Interested; 3: Slightly interested; 4: Not interested".

- Rural (Q13): Whether respondent district rural or not. Coded by interviewer [Yes=1, No=0].

- Friends protest (T905): Respondent was asked: "Did any of your friends or acquaintances participate in the protests against former president Ben Ali between December 17th, 2010 and January 14th, 2011?" [Yes=1, No=0].

- Civil society membership (Q5012, Q5013, Q5014, Q5016). Binary variable coded 1 if respondent answered in the affirmative to any of the above when asked: "Are you a member of? A charitable society (Q5012); A professional association/trade union (Q5013); A youth/cultural/sports organization (Q5014); Q5016: A local development association (Q5016)".

- Democratic commitment index (Q516): Composite index constructed using responses to three questions gauging commitment to democracy. A first (Q5162) asks respondents to respondent to the statement: "Democratic regimes are indecisive and full of problems. 1: I strongly agree; 2: I agree; 3: I disagree; 4: I strongly disagree". A second (Q5164) offers the statement: "A democratic system may have problems, yet it is better than other systems. 1: I strongly agree; 2: I agree; 3: I disagree; 4: I strongly disagree". A third (Q5167) offers the statement: "Democracy negatively affects social and ethical values in [Tunisia]. 1: I strongly agree; 2: I agree; 3: I disagree; 4: I strongly disagree". Items were rescaled so that higher values indicated stronger commitment to democracy. A scale of 0-4 was used, with missings recoded as the midpoint of the scal (2) in order to preserve sample size.

  The alpha on this index is relatively low (0.43). Excluding question Q5164 boosts the alpha to 0.61. Including another question (Q5163) alongside Q5162, Q5164, and Q5167, which offers respondents the statement "Democratic systems are not effective at maintaining order and stability", increases alpha to 0.65. Excluding Q5164 and using just Q5162, Q5163, and Q5167 gives an alpha of 0.78. For the purposes of the main analysis, I enter the index used by Doherty and Schraeder (2018) (a variant of an additive index used by Hoffman and Jamal 2014), in order properly to replicate their results and facilitate comparison when disaggregating by time of protest participation. The alternative indices with higher alpha give substantively identical results.

A.9.2 Proportional odds

The 'proportional odds" (or "underlying latent variable"/"parallel regressions") assumption is analogous to the assumptions underlying the threshold models in the formal modelling literature described in the main text. Here, the assumption

is that we can assume the existence of heterogeneous revolutionary thresholds in a population. While individual thresholds for participation might vary, an underlying continuous latent variable (e.g., revolutionary propensity) can be assumed to exist. Since the odds in an ordinal model can be interpreted as corresponding to the odds of exceeding a certain category, the categories of an ordinal scale have a threshold interpretation (Winship and Mare 1984). A violation of this assumption would mean that the odds of participating at each stage of the revolution were *not* proportional to each other, and would thus provide evidence against the conceptualization of protest in revolution as a unitary outcome of interest.

Another way to think about this is as a test of whether there is an underlying common "revolutionary propensity." This conceptualization understands protest participation as a threshold model. If enough people come out onto the street, those who are "less revolutionary" may be moved to act themselves. But if no such common propensity exists, which would be the case if the motivations for participation are not common but change over time, then we would expect to find evidence that this is assumption is violated.

A.10   Additional outputs and robustness checks

In this section I detail several supplementary analyses of protest event data using different ecological covariates of interest and provide full outputs for the survey analysis. I first provide the full regression outputs (Table A.6) for the VIIRS nighttime lights delegation-level measure of economic development with and without a time interaction. We can see that the results using this measure are substantively similar to those for the IDR measure. Including an interaction time significantly improves model fit as measured by AIC, BIC, and McFadden's R^2. Predictive margins and marginal effects of the nightlights measure over time and by stage are displayed in Figure A.5.

Table A.6: Discrete-time logistic regression of nighttime lights with and without time interaction, cluster robust standard errors

Variables	Model 1: Without time interaction	Model 2: With time interaction
Lagged protest	0.677**	0.347*
	(0.214)	(0.163)
Time		0.089***
		(0.015)
Nightlights	-0.099	-0.466**
	(0.058)	(0.163)
Nightlights*Time		0.021**
		(0.008)
Youth pop.	0.078	0.085
	(0.044)	(0.050)
Population (logged)	0.840***	0.888***
	(0.156)	(0.168)
Repression	0.225***	-0.020
	(0.049)	(0.066)
Constant	-13.970***	-15.808***
	(1.825)	(2.010)
Observations	5,894	5,894
AIC	1282	1245
BIC	1322	1299
McFadden's R^2	0.0720	0.101

For the additional event history analyses, I use the same model as in Table 1 but enter alternative ecological measures (see A.7). As a more basic measure of deprivation and development, I first test a variable measuring illiteracy rates at the delegation level (% illiterate population aged 10 or over). Given the apparent centrality of youth-led grievances relating to job opportunities, I also test a measure for delegation-level graduate unemployment (% unemployed with higher education certificate). Still others have pointed to the internet as playing a central role in the diffusion of revolutionary protest given the enhanced connectivity and new information transfer it affords (**?**). To test for the effects of internet connectivity on protest diffusion I include a delegation-level measure of internet usage (% households connected to Internet). These measures cannot all be included in the same model due to the high levels of collinearity betwen them (see A.8). They are also highly correlated with the IDR measure used in the main analysis. Including them all in a single model would thus not be meaningful and would be hindered by multicollinearity. I include them here simply in order to demonstrate that there is evidence of duration dependence when we use several different ecological covariates of interest, thereby providing
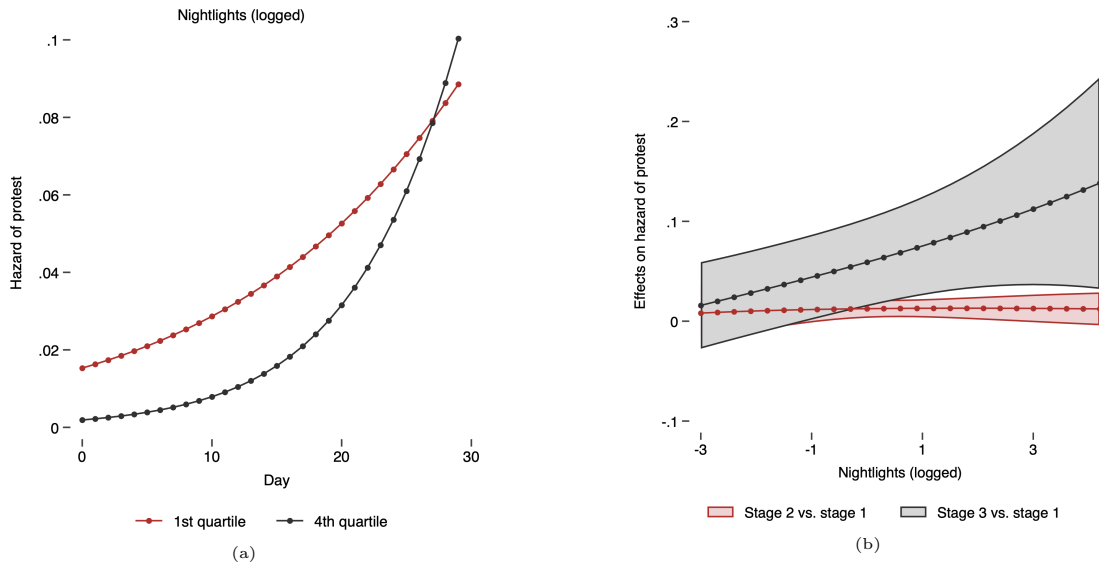
Figure A.5: (a) Predictive margins of nightlights over time, upper and lower quartiles; (b) Marginal effects of nightlights by stage of uprising, with 95% CIs

additional confidence in the robustness of the headline claim of significant duration dependence. As should be clear, the inclusion of a time interaction with each of these ecological covariates significantly improves model fit and points to significant time dependencies in the data. Figure A.6 displays predictive margins of each measure at upper and lower quartiles. Figure A.7 displays marginal effects for each stage of the uprising.

Table A.7: Supplementary event history analyses using alternative ecological covariates with and without time interactions

Variables	Model 1: Illit. rate w/o time interaction	Model 2: Illit. rate w/ time interaction	Model 3: Grad. unemp. rate w/o time interaction	Model 4: Grad. unemp. rate w/ time interaction	Model 5: Internet use w/o time interaction	Model 6: Internet use w/ time interaction
Lagged protest	0.627**	0.370**	0.633**	0.368*	0.584**	0.257
	(0.202)	(0.140)	(0.203)	(0.144)	(0.197)	(0.138)
Time		0.186***		0.219***		-0.054
		(0.029)		(0.036)		(0.032)
Ecological var.	0.018	0.095***	0.019*	0.104***	-0.001	-0.077***
	(0.010)	(0.020)	(0.009)	(0.021)	(0.007)	(0.020)
Ecological var.*Time		-0.004***		-0.005***		0.004***
		(0.001)		(0.001)		(0.001)
Youth pop.	0.089*	0.103*	0.059	0.062	0.063	0.083
	(0.045)	(0.049)	(0.042)	(0.050)	(0.046)	(0.051)
Population (logged)	0.848***	0.898***	0.845***	0.901***	0.738***	0.817***
	(0.151)	(0.160)	(0.149)	(0.158)	(0.146)	(0.157)
Repression	0.241***	-0.022	0.241***	-0.035	0.244***	-0.003
	(0.048)	(0.065)	(0.048)	(0.066)	(0.049)	(0.065)
Constant	-14.620***	-18.328***	-14.192***	-18.415***	-12.643***	-12.473***
	(1.952)	(2.172)	(1.761)	(1.967)	(1.660)	(1.825)
Observations	5,921	5,921	5,921	5,921	5,921	5,921
AIC	1288	1244	1287	1240	1291	1242
BIC	1328	1297	1327	1293	1331	1296
McFadden's R^2	0.0733	0.108	0.0741	0.111	0.0713	0.110

Robust standard errors in parentheses
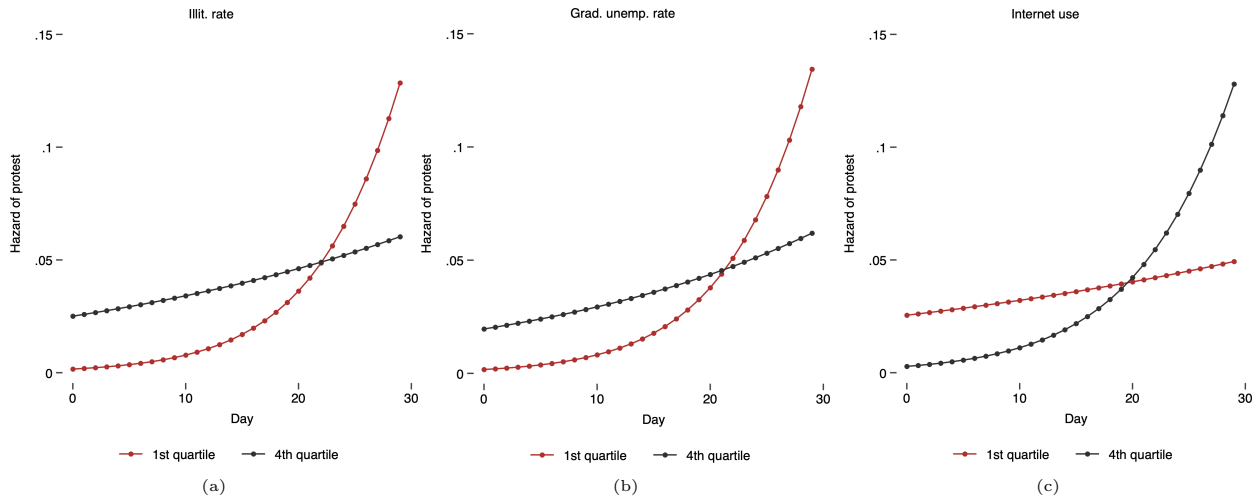*** $p<0.001$, ** $p<0.01$, * $p<0.05$

A15

Figure A.6: Predictive margins of alternative ecological measures over time, upper and lower quartiles. a): Illiteracy rate; b) Grad. unemp. rate; c): Internet usage
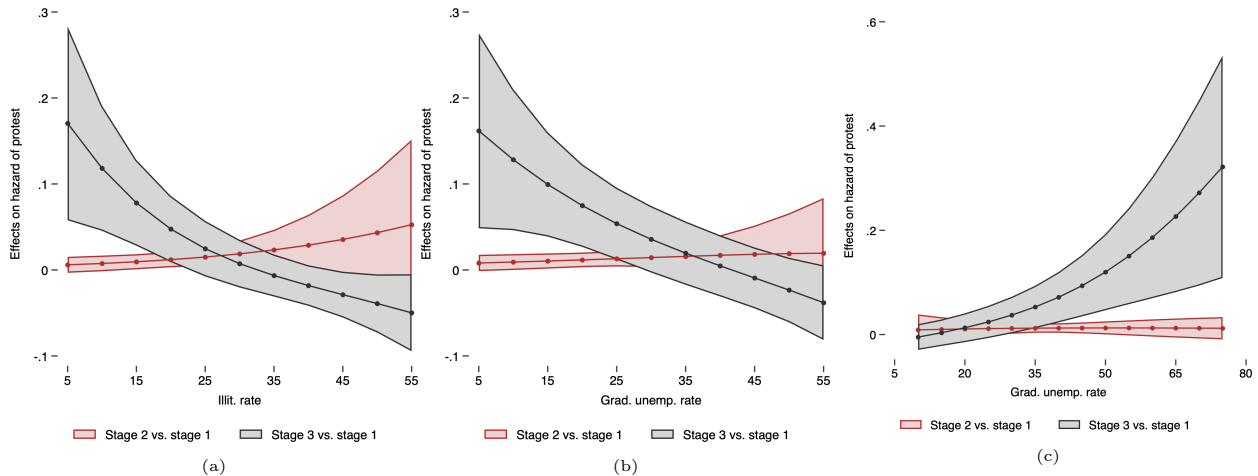

Figure A.7: Marginal effects of alternative ecological measures by stage of uprising, with 95% CIs. a): Illiteracy rate; b) Grad. unemp. rate; c): Internet usage

Table A.9 displays full results for the replicated analysis of Doherty and Schraeder (2018) alongside the results produced when using an ordinal outcome measure. Note that for ordinal models, cut points (not displayed here), are computed in place of a constant. Table A.10 displays the full results of the multinomial logistic regression with those who never protested used as the base category.

A.11 Methodological appendix—Mass mobilization for democracy events data

The following events were seeded by replicating Figure 4 in Hellmeier and Bernhard (2023), filtering to include only authoritarian polities (V-Dem polyarchy score $<.5$), and arranged in order of agreement between Carnegie events and expert coder evaluations. I consult online materials, principally Wikipedia and other published scholarship, to determine event details. The column "prp" records with a 1/0 whether the episode satisfies the twin characteristics of

Table A.8: Correlation matrix of alternative ecological variables

Variables	IDR	Illit. rate	Grad. unemp. rate	Internet use
IDR	1.000			
Illit. rate	-0.798	1.000		
Grad. unemp. rate 2014	-0.721	0.511	1.000	
Internet use	0.762	-0.883	-0.530	1.000

Table A.9: Logistic and ordinal logistic regression of protest participation with differently discretized outcome

Variables	Model 1: Logistic	Model 2: Ordinal logistic
Commitment to democracy	-0.052	0.083
	(0.155)	(0.158)
Income (logged)	-0.059	0.064
	(0.141)	(0.137)
Income (missing)	-0.062	-0.061
	(0.277)	(0.289)
Education	0.041	-0.034
	(0.087)	(0.082)
Age	-0.033**	0.031**
	(0.010)	(0.011)
Male	1.721***	-1.627***
	(0.283)	(0.293)
Employed	-0.038	-0.168
	(0.297)	(0.310)
Retired/Housewife	-0.859	0.757
	(0.459)	(0.508)
Student	0.219	-0.488
	(0.371)	(0.388)
Religiosity	0.453**	-0.440**
	(0.144)	(0.141)
Mosque attendance	-0.121	0.131
	(0.090)	(0.089)
Rural	-0.559*	0.518
	(0.274)	(0.293)
Interest in politics	0.284*	-0.315**
	(0.121)	(0.113)
Friend participated	2.342***	-2.516***
	(0.260)	(0.285)
Civil society membership	0.626	-0.390
	(0.344)	(0.323)
Constant	-3.406**	
	(1.227)	
Governorate indicators included	✓	✓
Observations	1,115	1,108

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

Table A.10: Multinomial logistic regression of protest participation with differently aggregated outcome

Variables	Model 1: Stage 1 vs. 0	Model 1: Stage 2 vs. 0	Model 1: Stage 3 vs. 0
Commitment to democracy	-0.261	-0.305	0.481*
	(0.222)	(0.282)	(0.235)
Income (logged)	-0.069	-0.356	0.061
	(0.208)	(0.308)	(0.220)
Income (missing)	0.300	-0.210	-0.103
	(0.405)	(0.669)	(0.397)
Education	0.025	0.370*	-0.057
	(0.127)	(0.167)	(0.122)
Age	-0.040*	-0.044	-0.029*
	(0.016)	(0.026)	(0.014)
Male	2.060***	2.381**	1.478***
	(0.478)	(0.822)	(0.438)
Employed	0.256	0.889	-0.575
	(0.456)	(0.698)	(0.378)
Retired/Housewife	-0.307	-13.476***	-1.569*
	(0.729)	(0.910)	(0.666)
Student	0.714	1.475*	-0.672
	(0.599)	(0.683)	(0.509)
Religiosity	0.612**	0.257	0.456*
	(0.218)	(0.301)	(0.212)
Mosque attendance	-0.204	-0.263	-0.080
	(0.130)	(0.234)	(0.135)
Rural	-0.509	-1.540*	-0.340
	(0.387)	(0.755)	(0.385)
Interest in politics	0.286	0.543*	0.310
	(0.163)	(0.261)	(0.170)
Friends participated	2.933***	2.453**	2.441***
	(0.448)	(0.753)	(0.380)
Civil society membership	0.745	0.741	0.385
	(0.501)	(0.558)	(0.437)
Constant	-3.420*	-20.689***	-21.917***
	(1.664)	(2.430)	(2.220)
Governorate indicators included	✓	✓	✓
Observations	1,108	1,108	1,108

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

A18

processual revolution as an episode that did not begin as revolutionary but saw: a) evolving demands as; b) new types of protestor joined the fold.

Table A.11: Sources of first event-days by delegation 17/12/2011–14/01/2011

country_name	year	protest_name	notes	prp
Belarus	2020	"Slipper uprising"	Classic case of protests in response to stolen elections by Lukashenko. See Mateo (2021) for event data.	
Armenia	2018	"#RejectSerzh" protests	Another case of protests in response to fears of authoritarian reversal under Sargasyan. Led by politician Nikol Pashinyan.	
Thailand	2020	Antigovernment protests	Began life as student-led protests with university-specific slogans, triggered by the dissolution of the FFP (political party). Continued after lockdown lifted in July 2020 with students advancing three demands: "dissolution of the House, ending intimidation of the people and drafting a new constitution." At this point, they explicitly stated they did not want to overthrow the monarchy. Then LGBT groups joined and added same-sex marriage to the demands. Then a Harry Potter-themed demonstration broke out in early August, explicitly demanding "amendment of increasing royal prerogative and reform of the lèse majesté law." Then on August 10 protest broke out with the slogan: ""We don't want reforms; we want revolution." By August 26, there were 10 demands and these were submitted to the House of Representatives. Protest meanwhile had grown to around 100,000. A plaque commemorating the People's PArty (which overthrew the monarchy in the early twentieth century) was installed by protesters–perceived as an explicit challenge to the King of Thailand. When news of delays to constitutional reform were announced, #RepublicofThailand started trending on Twitter.	1
Lebanon	2019	October Movement	Started in response to planned changes to tax regime and proposals to tax VoIP calls (such as on WhatsApp) as well as increases to petrol and tobacco tax. Developed into national protests against sectarian rule	1
Algeria	2019	Protests against "Le Pouvoir"	Protests began as very localized in Kabyle region and very small in size, but called for opposition to fifth term of Bouteflika. Ultimately united Berbers, Islamists, Leftists in calles for Bouteflika resignation–which they succeeded in getting later in 2019.	
Togo	2017	Term limit protests	Protests called in opposition to 50-year rule of Gnassingbé Eyadéma and Faure Gnassingbé that attracted hundreds of thousands onto the streets. Started life as explicitly in opposition to the government and calling for their removal.	

Nicaragua	2018	Social security reform protests	Protests began in response to social security reforms decreed by President Daniel Ortega. But protests occurred against backdrop of several years of protests, initially by pensioners but then by students and young people, against social security reform. Protests then took place over government handling of wildfires and the proposal of a deal with China to build a major canal across Nicaragua. Protests intensified and were heavily repressed, leading to outside intervention (by IAHCR) and the announcement of a "national dialogue." Arrests of protest leaders continued. Over time develops into explicitly anti-government protests, calling for removal and freedoms.	1
Malawi	2019	Election fraud protests	Protests against re-election of Mutharika amid claims of election fraud. In response to severe repression by the army, protesters called for resignation of the government, and end to police brutality, and removal of the head of the eleciton commission.	
Burma/Myanmar	2021	Coup protests	Protests that broke out in response to the coup staged by the leader of the armed forces in Burma/Myanmar.	
Iraq	2019	Iraq protests	Protests began in Basra as a highly localized protest against poor public services as well as unemployment. Then protest spread to Baghdad with other higher degree holders demanding employment opportunities. In response to repression and subsequent government mishandling of internal response, protests broke out nationwide against corruption and government incompetence. The protests combined both peaceful daytime demonstrations and nightime, youth-led, violent protests– the latter of which were forcefully repressed. Protestors began calling for resignation of Prime Minister Adil Abdul-Mahdi and chanted slogans against Iranian influence. PM resigns and protests continue to call for end to Iranian influence and electoral reform. SNTV electoral system is announced, which overcomes problems associated with lists. Over time, multiple new constituencies joined these protests, which culminated in calls to overthrow the Iraqi government.	1
Venezuela	2017	"Mother of All Marches" protests	Protests in response to attempted auto-golpe by the Supreme Tribunal of Justice. Massive potests spurred by opposition movements and began life as explicitly anti-government.	

Montenegro	2020	Religion law protests	Protests that broke out in response to a law that tried to strengthen hold of Montenegrin Orthodox Church by transferring property formerly held and controlled by the Serbian Orthodox Church. Seven out of ten Montenegrins still subscribe to the Serbian Church, which led to clashes on the streets and in parliament.	
Haiti	2018	PetroCaribe protests	Protests began in response to rising fuel prices but developed into explicitly anti-government protests. Social media references to PetroCaribe corruption began circulating, and protests continued. In the wake of repression and the release of a report on PetroCaribe corruption, large protests then broke out again in February, 2019. Protests began calling explicitly for President Moïse to step down.	1
Haiti	2020	"March for Life" protest	Against backdrop of previous protests, and a wave of violence and kidnappings that broke out, as well as calls by Moïse to make illegal e.g., burning tires on the street, protests break out again calling for Moïse to resign.	1
Haiti	2021	Presidential protests	Protests explicitly against Moïse began in January, 2021. Opposition leaders then allegedly attempted a coup, leading to a severe crackdown followed by calls by opposition leaders for a general strike. Protests continued against kidnappings and violence, led partly by mothers groups.	1
Cameroon	2018	Electoral protests	Protests in response to re-election of Paul Biya.	
Democratic Republic of the Congo	2017	Election delay protests	Protests and strikes in response to announcement of delay of elections.	
Mali	2020	Movement of June 5	Protests that began in wake of election irregularities, and triggered opposition-led protests calling for resignation of Ibrahim Boubacar Keïta, which resulted in eventual coup.	
Cameroon	2017	Ambazonia protests	Protests broke out in response to the hiring of Francophone judges in the Anglophone north of Cameroon. This developed into region-wide protests opposing the government of Paul Biya and calling for Ambazonian independence.	1
Zimbabwe	2017	Mugabe protest	Protests against ongoing corruption and poor economic performance broke out in Zimbabwe. Developed into calls for the resignation of President Robert Mugabe. The initial protests followed a viral video by Pastor Evan Mawarire calling for Zimbabweans to speak up. Protests were eventually joined by traditionally pro-Mugabe groups like the veterans association (the ZNLWVA). The ZNLWVA was then central in subsequent protests explicitly calling for Mugabe to step down.	1
Sudan	2019	National Congress Party protests	Protests began in Atbara in respond to cuts in bread subsidies. These protests then escalated into nationwide protests against President al-Bashir.	1

Ethiopia	2018	Ethnic violence protests, Political prisoner protests	Oromo-linked ethnic violence leads to mass protests against government. This came after a series of protests that ousted former leader Abiy and ushered in democratic reforms.	
Bolivia	2020	2020 election protests	Protests in response to election delays amid ongoing political crisis.	
Mali	2019	Ethnic violence protest	Protests in wake of deadly ethnic violence in Fulani community. United a broad coalition of religious leaders and civil society groups to protest against the government of Ibrahim Boubacar Keita. Led to sustained calls for protest and ultimately a coup.	1
Chad	2021	Military rule protests, Re-election protests	Protests in response to re-election of Idriss Déby who had ruled the country for over twenty years.	
Kyrgyzstan	2020	Electoral protests, Internet freedom protests	Protests stimulated allegations of vote-buying in the October elections. Resulted in the ultimate resignation of President Jeenbekov a week later.	
Albania	2019	Local elections protests	Protests motivated by allegations of vote-rigging by prime minister Edi Rama.	
Lebanon	2020	Prison protest	Protests in Lebanese protests over conditions in context of Covid-19 pandemic.	
Sudan	2020	"Million-man March", Darfur insecurity protests	Protests that are part of a long wave of protest in Sudan that began in Atbara over bread subsidies.	1
Zimbabwe	2018	Electoral protest	Protests in response to results of the 2018 election and the regime of Emmerson Mnangagwa.	

References

Cambon, J., Hernangómez, D., Belanger, C., & Possenriede, D. (2021). *tidygeocoder: An R package for geocoding.* Journal of Open Source Software, 6(65), 3544.

Doherty, D., & Schraeder, P. J. (2018). *Social Signals and Participation in the Tunisian Revolution.* Journal of Politics.

Elvidge, C. D., Baugh, K. E., Zhizhin, M., & Hsu, F.-C. (2013). *Why VIIRS Data Are Superior to DMSP for Mapping Nighttime Lights.* Proceedings of the Asia-Pacific Advanced Network, 35, 62.

Hellmeier, S., & Bernhard, M. (2023). *Regime Transformation From Below: Mobilization for Democracy and Autocracy From 1900 to 2021.* Comparative Political Studies, 00104140231152793.

Hoffman, M., & Jamal, A. (2014). *Religion in the Arab Spring: Between Two Competing Narratives.* Journal of Politics, 76(3), 593-606.

Howard, P. N., & Hussain, M. M. (2014). *Democracy's Fourth Wave? Digital Media and the Arab Spring.* Oxford University Press.

Kondo, K. (2017). *Introduction to Spatial Econometric Analysis: Creating Spatially Lagged Variables in Stata.* Statistical Software Components S458105 Boston College Department of Economics.

Winship, C., & Mare, R. D. (1984). *Regression Models with Ordinal Variables.* American Sociological Review, 49(4), 512.

Wu, K., & Wang, X. (2019). *Aligning Pixel Values of DMSP and VIIRS Nighttime Light Images to Evaluate Urban Dynamics.* Remote Sensing, 11(12), 1463.