Narrative

Burak Tan

09 12 2019

Brief Substantive Background and Goal

The Republic of Turkey had its 25th general elections on June 7th, 2015, in which no single party won the necessary number of seats for the parliamentary majority necessary for the formation of a single party government. Following the failed attempts to form a coalition government, the country went for a swift re-election on November 1st, 2015. In this second election, the incumbent Justice and Development Party (AKP) increased its vote share by 8% and gained 59 additional seats at the parliament (10.7% of the total number of seats). The steep increase in votes over the short span of time coinciding with the resurgance of the Kurdish civil war and pervasive security operations in the southeastern regions raised the question of fairness for the elections.

This project evaluates whether there is a relation between the local disruptions of state violence and the increase in AKP votes. For this evaluation, I scraped the website of the Turkish Supreme Election Council and calculated the percentage increase in votes on election district level. Then, I generalized this data to the level of geographical region, using the NUTS classification of the European Union and used UCDP Georeferenced Event Dataset (battle related death) as an indicator of civil war's impact on the region.

I hope to shed further light on resurgence of the Kurdish Civil War in Turkey between these months, by determining whether there is a significant relation between the districts that were affected from the military operations and the cities in which the election results changed.

Collecting, Cleaning, Pre-Processing Data

There are three main sources for the data used in this project: First, the website of the Turkish Supreme Election Council; second, the website of the Eurostat; and third, the website of the UCDP.

1. Election Results

The website of the Turkish Supreme Election Council does not provide statistics for district level results in the general elections. Results for each general election are reported on different pages, with detailed options to select via drop down menus. For each election result (June 2015 and November 2015), one needs to select:

1. Confirmed Election Results ("Kesin Seçim Sonuçları" through the menu on the left)
2. District Level Results ("Seçim Çevreleri Sonuçları" through the menu on the middle)
3. District name (dropdown menu)

This process produces a PDF file with all the results for one election district out of 85, in a convoluted page set up. After trying to scrape the website leveraging the JavaScript that runs the menu options and dropdown lists, I decided to pursue an alternative strategy.

Conveniently, the PDFs produced are pulled from the servers of the Council through calling a URL that take the form of:

http://www.ysk.gov.tr/doc/dosyalar/docs/Milletvekili/7Haziran2015/KesinSecimSonuclari/Cevre/Adana.pdf

Which can be broken down and translated as:

- "http://www.ysk.gov.tr/doc/dosyalar/docs/" +
- "/Election Type" +
- "/Election Date" +
- "/Certified Election Results" +
- "/District Level" +
- "District Name" +
- ".pdf"

This structure enables the leveraging of the URLs themselves, rather than going through the complete HTML code of the Council website.

Hence, the scraping of the AKP vote counts on these PDFs was achieved in two steps:

1. Creating a list of URLs by substituting the "District Name" with a vector of election districts
2. Mapping a scraper function that reads the PDF and pulls the AKP vote count as a numeric value on the list of URLs.

The First Challenge: Inconsistent Bureaucrats!

A brief survey of a couple PDFs quickly revealed that Turkish bureaucrats working for the Election Council are not big on consistency! The results for the two elections were reported in similarly formatted PDFs, yet the inconsistencies were enough for the scraper function to stop working. Specifically, the vote counts reported as "abc,def" and "abc.def"; and the different order of the political parties reported on the PDFs made two scraper functions necessary.

The Second Challenge: Non-Latin Characters of the Turkish Alphabet

The non-Latin characters of the Turkish Alphabet that are frequently used in the titles of the election district names made it impossible to use the list of districts available as a PDF on the Council website. The challenge was not to change each character to another one, but to change each character to the one that was used in the specific case of the election district's URL. Particularly the letters "ı" and "i" capitalized as "I" and "İ" respectively found different uses in the title of the districts. Hence, a vector was created for the most part manually, by copy-pasting the list of districts and fixing each district's title where the scraper function broke. This required going back to the code that creates the URLs, each time mapping did not work, by subsetting the URL list ten at a time for each attempts and checking whether the function creates a numeric value as result.

2. NUTS District and Region Codes

As a candidate state, Turkish Republic is divided into regions by the European Union, and the information is both widely applied and consistent with the country's geography. Yet, after trying the pdf_text and read_excel functions on various formats Eurostat website provides, and a myriad of options to read excel files, I came to admit that the amount of formatting of the excel made it impossible to accurately read it.

My solution was to create a new dataframe copy-pasting the list of districts and list of regions from the excel document, and adding the election district that were not just cities. There were three cases that NUTS (as a city-based list of regions) did not include, these were Ankara, Izmir, and Istanbul. To ensure that the dataframe is consistent with the election-district level results produced in the first dataset, I added these manually. Producing the dataframe manually also enabled consistency across the rest of the election district titles, which then enabled the merging of datasets.

3. UCDP Data on Battle Related Deaths

I decided to use the UCDP Georeferenced Event Dataset (GED) Global version 19.1 as the independent variable referring to the pervasiveness of the civil war in these districts. Whereas it is abundantly clear that on the regional level the civil war affected Southeastern and Eastern Turkey (the predominantly Kurdish regions), district level impact was not so obvious. While the pervasiveness of the effects cannot be limited to the amount of battle related deaths, the georeferenced dataset UCDP provides is a good estimation of the violence and dislocation endured.

As UCDP provides remarkably clean data, pre-processing was significantly easier than the case of the Election Council. Once the agents (Government of Turkey) and the dates (June 6 - November 1) were determined, I filtered the dataset and formatted the location information (column adm_1 which provides the title of the city) to fit with the coding of the district vector in use. While there are several useful tools such as the stri_trans_general and chartr functions, once again the inconsistent capitalization and Latin-Turkish translations of the district names made it impossible to get rid of all the inconsistencies. Hence, I included a piped str_replace_all operation, with the specific changes required.

The unique case of Istanbul, where deaths occurred and were reported by the UCDP on a city level, was addressed through allocating the amount of casualties equally to the three election districts. Once all the districts where deaths occurred was secured in a tidy dataset that is remarkably smaller than the UCDP data, I added the rest of the districts with the rbind function, so that this dataframe can be merged into the final dataset.

Analysis and Visualization

Both the change in AKP votes derived from the PDFs and the information on the battle related deaths are calculated as percentage change in the final data. The reasoning behind this choice is fairly simple: for the first case, I am interested in the change in the amount of votes based on what the political party can achieve on that district. For example, an increase of 50,000 votes indicates very different degrees of success in Igdir, where the total amount of votes counted was 94,909, compared to Istanbul1 (one third of the city), with 3,036,754 counted votes in June 2015 elections. In the case of the battle related deaths, I am similarly interested in the share of that district in terms of the civil war's toil, rather than the exact number of deaths. While the battle related deaths are not a complete measure whatsoever for this burden, where does deaths occurred gives a good starting point in understanding the geographical distribution of the violence.

Visualization of the Data

For the visualization of the final data, I chose to secure the dataset in an item and and amend it ever so slightly through mutation of additional column for the purpose of the specific visualization.

- Plot_0 is a descriptive chart that demonstrates the Percentage Change in AKP Votes by District, showing that the votes of the ruling party increased in every single region, albeit in different ratios.
- Plot_1 demonstrates the Change in Votes by Region, indicating that three regions stand out in terms of the percentage increase in AKP votes. These are Eastern Anatolia, Northeastern Anatolia, and Southeastern Anatolia.
- Plot_2 shows the Distribution of Wartime Deaths by Region, between June-November 2015, proving the assumption that the toil of the civil war was predominantly shared by the Southeastern and Eastern Turkey.
- Plot_3 visualizes the Increase in AKP Votes by Region's Share of Wartime Casualties, where a pattern begins to emerge between the two variables.
- Plot_4 takes the information in the previous visualization to the election district level and shows the Share of Deaths by District and Percentage Increase in AKP Votes. Indicating the pattern with a smoothed linear model line.

- Plot_5 provides the same information with the Plot_4, with the x-axis logged, to demonstrate the results more clearly.

Regression Results

The regression results use three basic linear models, using the Percentage Increase in AKP Votes as the dependent variable.

First model considers District's Share of Deaths as the independent variable, and finds a strong relationship (0.906) even though the degree of confidence is limited to $p < .1$. This lack of parsimony may be interpreted as a result of the small N, which is limited to the number of districts, that is to say, 85.

The second model incorporates Region's Share of Deaths to the former as an independent variable and finds a very strong relationship between the two (2.393) with a reasonable degree of confidence ($p < .05$). This confirms the hypothesis that there was a relationship between the increase in votes AKP enjoyed between the June and November elections of 2015 and the civil war, not only as a result of the effects civil war had on the country as a whole, but at the level of geographic region. AKP increased its percentage share of the votes significantly more in the regions that were most strongly affected by the civil war.

The third model looks at the districts within regions showing homogeneity, with the exception of the North Eastern Anatolia, where the city of Igdir, which is an election district of its own has an increase of 1.799 (roughly 180%) meaning the party practically tripled its vote share. While the exact reasons of this are unclear, the very small number of voters in the district certainly contributes to the skewed results.

Future work

The study demonstrated that there is a significant relation between the geography that was affected from the military operations and the cities in which the election results changed for the favor of the ruling Justice and Development Party. While this relation is weaker at the level of electoral district, it is clear at the level of regions.

These findings consolidate the allegations circulated in the media against the AKP, considering the freedom and the fairness of the November 2015 elections within the context of extreme state violence. Further work may test these results through additional data shedding more light on the pervasive influence of the civil war in these regions, or through the addition of various control variables such as the economic status of the country in June and November, or demographic information. Political scientists interested in Turkish Politics and in the Middle East more broadly may build up on these results to ascertain the intentionality of this outcome through an ethnographic study of regions that were affected by the violence at the time of the second election, and the party cadres themselves.