Twitter and the European Hyperagora: What can the Twittersphere Tell us about Political Deliberation and Opinions in Europe?

M. Callaghan and V. Niberg

December 11, 2015

Contents

Introduction

Twitter is an online social network that allows users to broadcast short posts known as Tweets. Since its launch in 2006, the platform has increasingly been used for everyday communication as well as for political debates, crisis communication, marketing, and cultural participation (Weller et al. 2013). The public-debt crisis in Europe is widely discussed across Europe and presents an interesting point in time to investigate whether European issues are discussed in a common European public sphere. This project looks at data from the communication platform Twitter. It specifically looks at the reaction in the Twittersphere to the negotiation between the Troika and Greece leading up to the signing of the three memorandums.

Research Question

The following questions were investigated:

What can twitter tell us about pan-European reactions to the European governance of the public-debt crisis in Greece?

- What can variation across time and space in the volume of Tweets regarding the euro crisis tell us about popular engagement with the issues?
- What can the content of Tweets related to the crisis tell us about the spread of public opinion on the handling of the crisis in Greece between and within countries?

The answer to these questions could potentially add to the literature on the emergence of a European public sphere.

Literature Review

On Twitter Research

The body of twitter research has grown steadily in recent years (for a comprehensive analysis and typology of twitter research up to 2013, see Zimmer and Proferes 2014). Some of the findings relevant to our research design are discussed below.

Twitter is a source of meaningful information about engagement with and opinions about political topics. Twitter is also used as a platform for political deliberation. In a recent study on Tweets mentioning parties or politicians before the 2009 German federal election, Tumasjan et al. found that "Twitter is not just used to spread political opinions, but also to discuss these opinions with other users" (Tumasjan et al. 2010, 183). Furthermore, specific patterns of twitter usage have been identifed that correspond with high-profile political events. Hughes and Palen found that, compared to general Twitter usage, more broadcast-based information sharing activities take place (Hughes and Palen 2009, 259) during events. Moreover, Tumasjan et al. found that it was possible to extact meaningful information about political opinions from both the volumes and the content of these Tweets: "the mere number of Tweets reflects voter preferences and came close to traditional election polls" (Tumasjan et al. 2010, 183).

Twitter gives information on the location of Tweets and users, which must be carefully interpreted. Devin Gaffney points out methodological problems with using the given location of twitter users - "in many cases user-entered profile locations differ from the physical locations users are actually Tweeting from" (Graham, Hale, and Gaffney,Devin 2014, 1) which must be considered when interepreteing results.

Though the field of Sentiment Analysis (SA) is perhaps most developed in the business world (Zimmer and Proferes 2014, 250), an increasing body of literature has developed, focused on retrieving information about political opinions from the Twittersphere. Though Tumasjan's results have come under scrutiny (see Jungherr, Jürgens, and Schoen 2012), the authors found that "the sentiment of Twiter messages closely corresponded to political programs, candidate profiles, and evidence from the media coverage of the campaign trail" (Tumasjan et al. 2010, 183).

Grimmer provides an overview of recent developments in SA in political science, noting how "automated content methods can make possible the previously impossible in political science: the systematic analysis of large-scale text collections without massive funding support" (Grimmer and Stewart 2013, 2). He advises caution, however, about the utility of SA in predictive models: "The goal of building text models is therefore different than model building to make causal inferences. […] Emphasis in evaluations should be placed on helping researchers to assign documents into predetermined categories, discover new and useful categorizytion schemes for texts, or in measuring theoretically relevant quantities from large collections of text." (Grimmer and Stewart 2013, 4).

Due to the enourmous amount of text available, Pak and Paroubek identify that "microblogging web-sites are rich sources of data for opinion mining and sentiment analysis" (Pak and Paroubek 2010, 1320). The multilingual nature of Tweets across Europe presents some difficulties, but is the subject of a growing body of research: "Noisy social media, such as Twitter, are especially interesting for sentiment analysis (SA) […] given the amount of data and their popularity in different countries, where users simultaneously publish opinions about the same topic in different languages" (Vilares, Alonso, and Gómez-Rodriguez 2015, 2). Balahur and Turchi are confident about the ability of Statistical Machine Translation (SMT) to provide a basis for consistently applied SA across languages (Balahur and Turchi 2012, 58). Other approaches include using emoticons to train models that assign sentiment to a multilingual text corpus (Narr, Hulfenhaus, and Albayrak 2012).

Finally, some studies discuss ethical aspects of twitter research. For example, concerns about creating a permanent archive of Tweets have been voiced. These concerns included whether "such archive was aligned with users' privacy expectations" (Zimmer and Proferes 2014, 258; Zimmer 2010).

On Awareness and Public Opinion across Europe on the Governance of the Public-Debt Crisis in Greece

Academic research on the emergence of a European public sphere is not a recent phenomenon (Risse 2003, 1). Hitherto, however, research has been characterized as rather normative, as the "research community has been […] interested in producing policy recommendations for public sphere-building" (Trenz 2015, 234). Recent studies, on the other hand, seem to put emphasis on an empirical grounding of the debate (Trenz 2015; Drewski, Gerhards, and others 2015). This development is being mirrored in research on the public debate across Europe on the euro crisis. It has been suggested that "there is an emerging demos in the European polity and it has been strengthened during the euro crisis" (Risse 2014, 1213). When testing this hypothesis empirically, though, by looking at newspaper editorials in Spain and Germany, Drewski found that there were significant differences along national instead of ideological lines in the discussion of the Euro crisis (Drewski, Gerhards, and others 2015, 5).

Max Hänska and Stefan Bauchowitz in a recent LSE blog entry track twitter activity during the negotiations leading up to the third Greek bailout agreement. (Haenska and Bauchowitz 2015) According to their findings, Tweets synchronised around key mini-events throughout the negotiations, with peaks and troughs mirrored across national twitter-spheres. These results suggest that popular engagement with the issue converges across Europe.

They further looked at instances of Tweets containing #ThisIsACoup, representing a particular opinion on the agreement. They then showed that the spread of #ThisIsACoup was not reflected in the studied countries equally. This indicated a divergence of public opinion along national lines.

Data Sources

For our investigation in the European public discourse on the Euro crisis, two datasets were required. The first was the corpus of Tweets relating to the Greek debt crisis and the measures taken to manage the crisis by

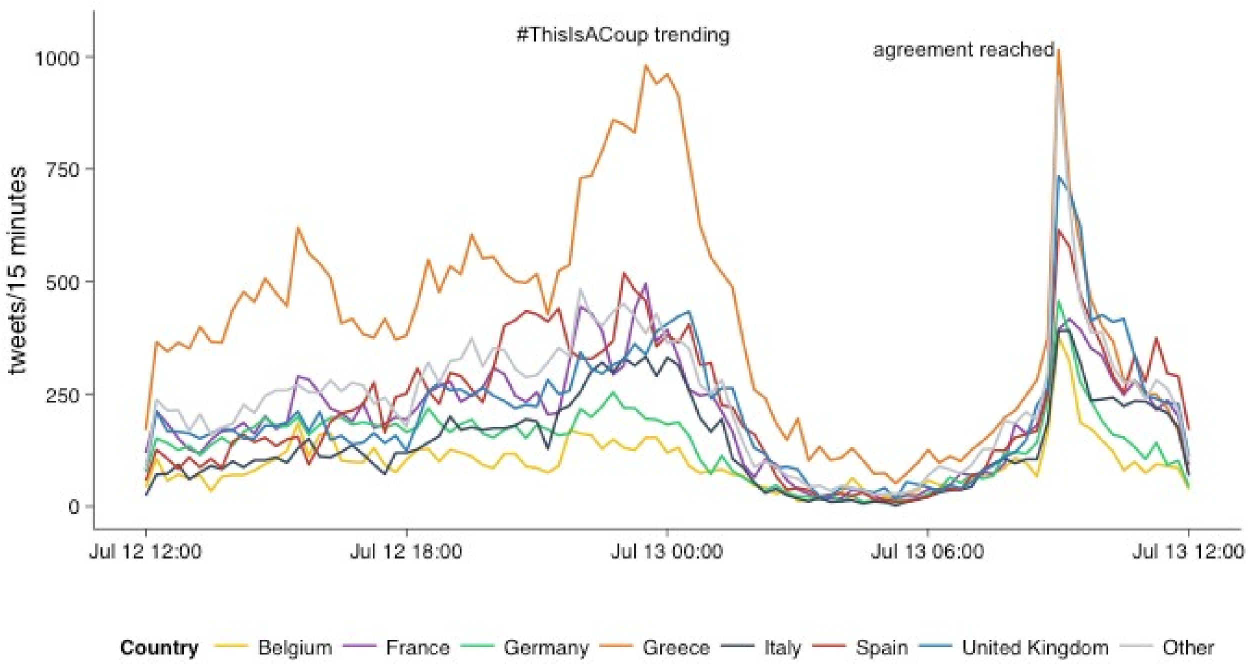

Figure 1: Tweet volumes by country on 12-13 July 2015 in European countries (source Haenska and Bauchowitz 2015)

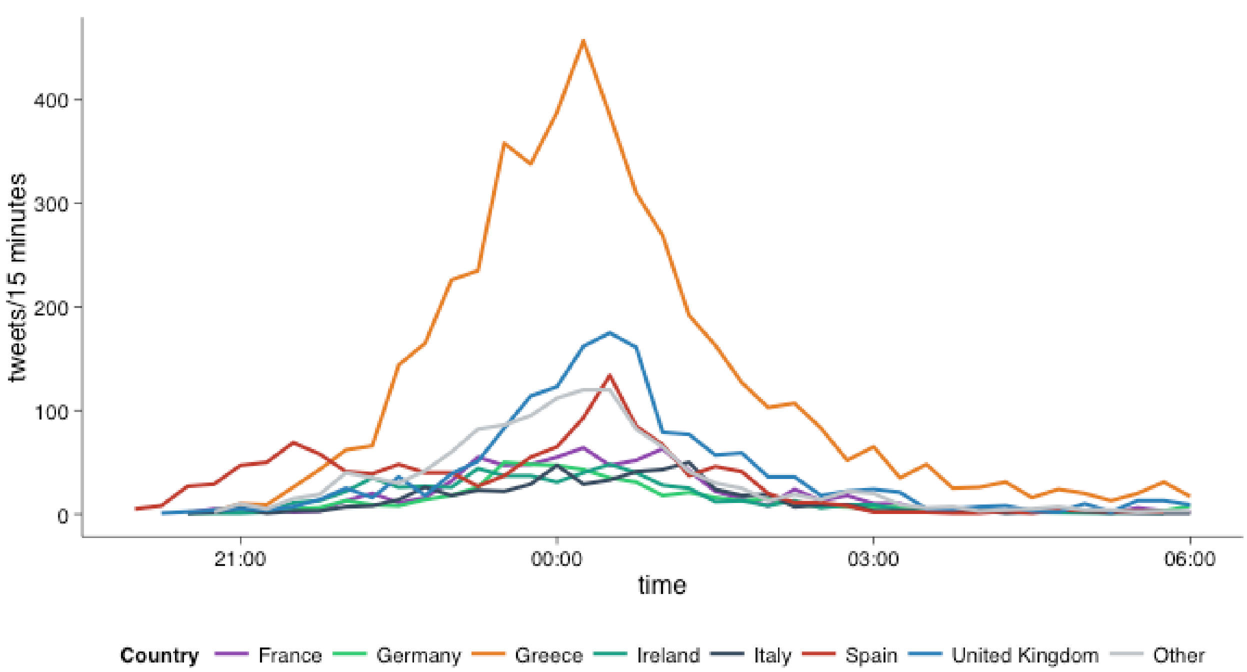

Figure 2: Number of Tweets containing the #thisisacoup hashtag on 12-13 July 2015 (source Haenska and Bauchowitz 2015)

European institutions. The second was information about the users whose Tweets form the body of that corpus.

Zimmer and Proferes identify the Library of Congress' decision to place every Tweet since Twitter's inception in 2006 into an archive as validating "the research importance of twitter" (Zimmer and Proferes 2014, 251). Despite this announcement occuring in 2010, five years later, the archive is still not open to researchers (Scola 2015). Since late 2014, the whole corpus of twitter data has been searchable online (Metz 2015). Programmatic access to this archive is, however, more restricted. Twitter's public search API "is not complete index of all Tweets, but instead an index of recent Tweets. At the moment that index includes between 6-9 days of Tweets." ("The Search API" 2015). Twitter sells access to historical Tweets through an API provided by its "enterprise API platform" GNIP (Tornes 2015). This paper adapts a publicly available program written in Java which scrapes results from Twitter's online search page (Henrique 2015).

Data Gathering

Due to the large amount of data we process, we ran the data gathering and cleaning in the background on a server using the prefix setsid.

Tweets

We used a modified version of GetOldTweets (Henrique 2015), a Java program that scrapes data from twitter search. The file getting_tweets/input.txt contains a list of search terms related to the Greek crisis in three periods, each comprising some weeks before and after the negotiation and signing of the memoranda. The search terms were collected using an adapted form of snowball sampling (Biernacki and Waldorf 1981), searching an initial list and recursively adding related terms found in the results. By running

```
sudo setsid ./compile_run.sh ../getting_tweets/input.txt
```

from the GetOldTweets folder, we ran through each search term and each period, searched twitter, and saved the results as a txt file in the data folder. After an initial assessment of the results, we refined our search terms and ran GetOldTweets again with /getting_tweets/input2.txt. A third file (getting_tweets/input3.txt) aims to return a time-inpedependent list of tweets in order to control for the growth of Twitter over time.

We end up with a long list of files in the data/GOToutput folder, which in the data cleaning process will be merged into one corpus file.

Users

We found the unique users in our corpus of tweets and used the TwitteR package (Gentry 2015) to gather richer data about each user. TwitteR uses the twitter API and gives the opportunity to collect all information twitter has abou the user. Where a users's last tweet was geocoded, we took the latitude and longtitude. We end up with the file data/user_info.csv

Many users do not geotag their tweets, instead stating their location, and we used APIs from MapQuest and Google to geocode user-reported location, giving us the file places.csv.

Merging & Cleaning

The txt files containing the tweets for each query and period are merged into a corpus file. This corpus file was merged with the user_info file, which in turn was merged with the places file. We end up with a large file containing tweets for our queries in each period with elaborete user information.

Some of the queries we defined returned irrelevant data, due to their ambiguity. We identified these by selecting random tweets from the search queries, reading the tweets, and checking for relevance to the topic. For example, the query “bailout”, although certainly relevant for our topic, was insufficiently precise and returned a lot of data about the banking bailouts, especially in the 2010 period. The following list summarizes the queries which we excluded.

- athens
- bailout
- 2-pac
- 3-pac

Translation

In a next step we identified the language of every tweet (Hornik et al. 2015) and translated those tweets written in European languages other than English into English (Lucas and Tingley 2014). Danish and Romanian were excluded from the list because these languages were not adequately identified by the program.

Final Dataset

A dataframe ‘merged_corpus’ containing all of the above information was pulled together. We filtered all tweets from Europe.

Methodology

Volumes of topic-relevant Tweets were mapped across space and time, to analyse the distribution of topic-awareness and its relation to political developments in responses to the crisis. The distribution of hashtags that clearly represent an opinion on the response to the crisis (e.g. ‘#ThisIsaCoup’, ‘#ThisIsNotaCoup’ *inter alia*) were similarly mapped in order to approximate the distribution of opinion within and between countries over time. The paper attempts a sentiment analysis of Tweets expressing opinions about the agreed bailout deals using machine translation to translate all texts into English before performing sentiment analysis. Analysis was then carried out using unigrams to indicate polarity through comparison with a lexicon. Following Grimmer, we assumed “documents are a *bag of words*, where order does not inform our analyses” as “In practice, for common tasks like measuring sentiment, topic modeling, or search, *n-grams* (combinations of words rather than individual words) do little to enhance performance” (Grimmer and Stewart 2013, 6).

To carry out the sentiment analysis process, we use the R package tm.plugin.sentiment (Annau 2014) to compare the words in our translated corpus of tweets with the General Enquirer lecixon (Theussl, Hofmarcher, and Hornik 2015). Each tweets is given a positive score and a negative score according to the number of positive and negative words found in the tweet. We calculate an overall sentiment score by subtracting the negative score from the positive score. Low scores therefore indicate negative sentiment; high scores indicate positive sentiment.

Based on the results of sentiment analysis carried out and analysis of volumes of tweets and users using opinion-signifying keywords, the paper gives an indication of the scale of dialogue, consensus and disagreement across and within countries in the European Twittersphere.

Analysis

Descriptive Statistics

The word cloud gives us an overall picture of the words used in the collected tweets connected to the European public-debt crisis.[1]

[1] to a certain degree in this case a word cloud is redundant, because it mainly returns our query turns. However, it also indicates the prevalence of specific terms, and further, which other terms are mentioned regularly in those tweets.

Figure 3: Word Cloud of a random sample of 200 tweets in our corpus

The following tables describe variation over space and time of the entire population of tweets.

Variation Across Space

This bar chart shows the distribution of all tweets over countries filtering those with a significant amount of tweets. This group seems to be a mix of high twitter user numbers and affectedness/involvement/relation to the events in Greece. Most probably reflecting high user numbers in the US, the US seems to be the origin of most tweets regarding the European sovereign-debt crisis. However, when controlling for period, the US resembles an atypical case. While for most other countries, tweet number on the topic increase over the years, they decrease substantially in the US.

Variation Across Time

The results show that twitter coverage of the Greek bailouts has increased heavily over time. When looking at the development of different queries over time, an increase was found for all queries except for for imf+greece, which decreased in 2012 and increased in 2015 again. Since the shown results are not normalized, it is difficult to say how much of the growth of the population of tweets can be attributed to an increase in twitter usage or to an increase of interest in the topic. Further, some of our queries only became topical in the later periods or even the last periods.

Variation Across Space and Time

Queries

The following table summarizes the collected data by queries. It shows absolute numbers and relative distributions of specific query returns.

query	n	percent
#aGreekment	3045	2.15
athens+berlin	478	0.34

query	n	percent
austerity+greece	2833	2.00
economic+adjustment	55	0.04
eucrisis	117	0.08
eurocrisis	2582	1.82
eurogroup	1421	1.00
eurogroup+agreement	312	0.22
eurosummit	8969	6.33
euro+summit	10211	7.20
germany+greece	3861	2.72
greece+crisis	3591	2.53
greece+reforms	4297	3.03
grexit	5515	3.89
imf+greece	2451	1.73
memorandum+greece	493	0.35
merkel+greece	5443	3.84
mou+greece	451	0.32
#nai	4055	2.86
#nai+greece	144	0.10
notmyeurope	255	0.18
#oxi	1756	1.24
rescue packages	41	0.03
schäuble+greece	2465	1.74
syriza	2873	2.03
tax+evasion+greece	298	0.21
#thisisacoup	6802	4.80
#thisisnotacoup	303	0.21
tsipras	6522	4.60
varoufakis	13642	9.62
bailout+eurozone	4482	3.16
bailout+greece	6342	4.47
bailout+greek	1747	1.23
efsm+greece	390	0.28
efsm+greek	74	0.05
ems+greece	14	0.01
esm+greek	340	0.24
greece+debt	4905	3.46
greek+crisis	20573	14.51
greek+debt	4676	3.30
greek+reforms	2774	1.96
memorandum+greek	170	0.12
six-pack+greece	2	0.00
six-pack+greek	3	0.00

The results that our queries returned differ substanially in size. While some queries (e.g. greek+crisis) return over 16000 tweets, others (e.g. economic+adjustment, rescue+packages) return only around 30 tweets in the specified time periods. We therefore think it is for analytical reasons useful to differentiate between high-return and low-return queries. The following charts report the developments over time for the high return rates as this group creates more intelligle results.

n the 2010 period, “greek+crisis” and “bailout+greece” are the most frequently used query. “Bailout” usage in tweets peaks around April 23rd when Greek prime minister Papandreo formally requested an international bailout for Greece. In the following day, “greek+crisis” is used at its highest level, with nearly 125 occurances

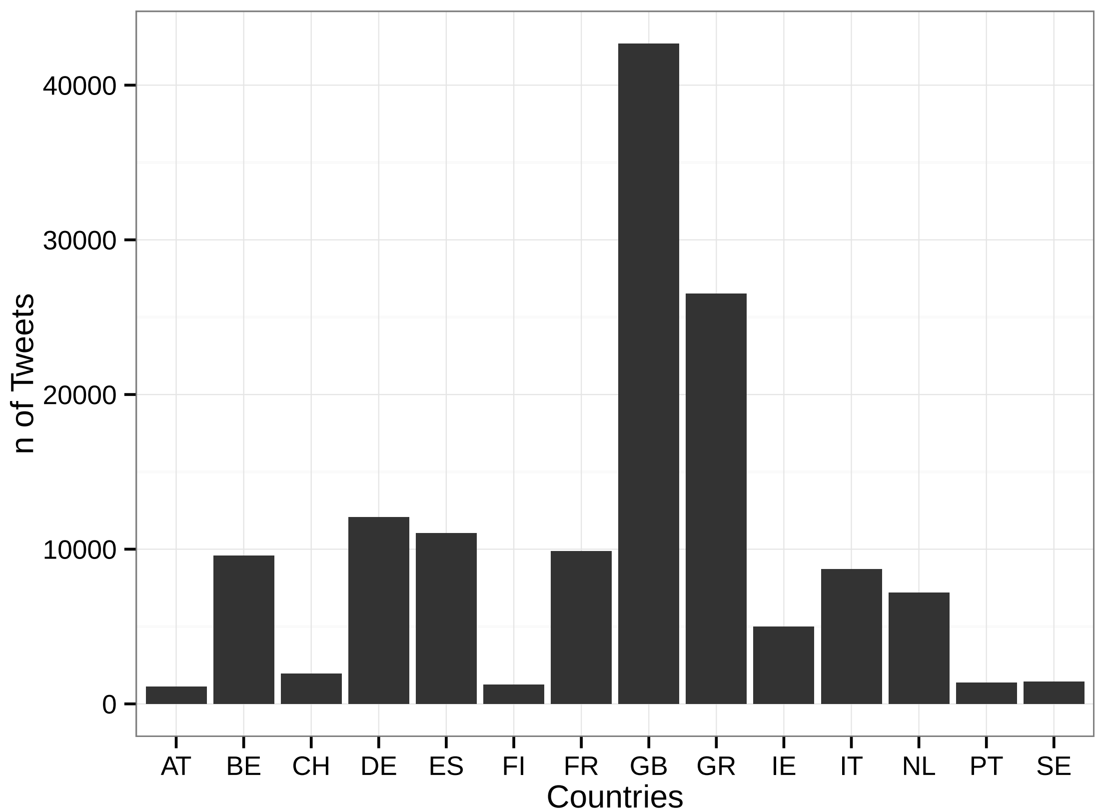

Figure 4: Geographical Distribution of Tweets

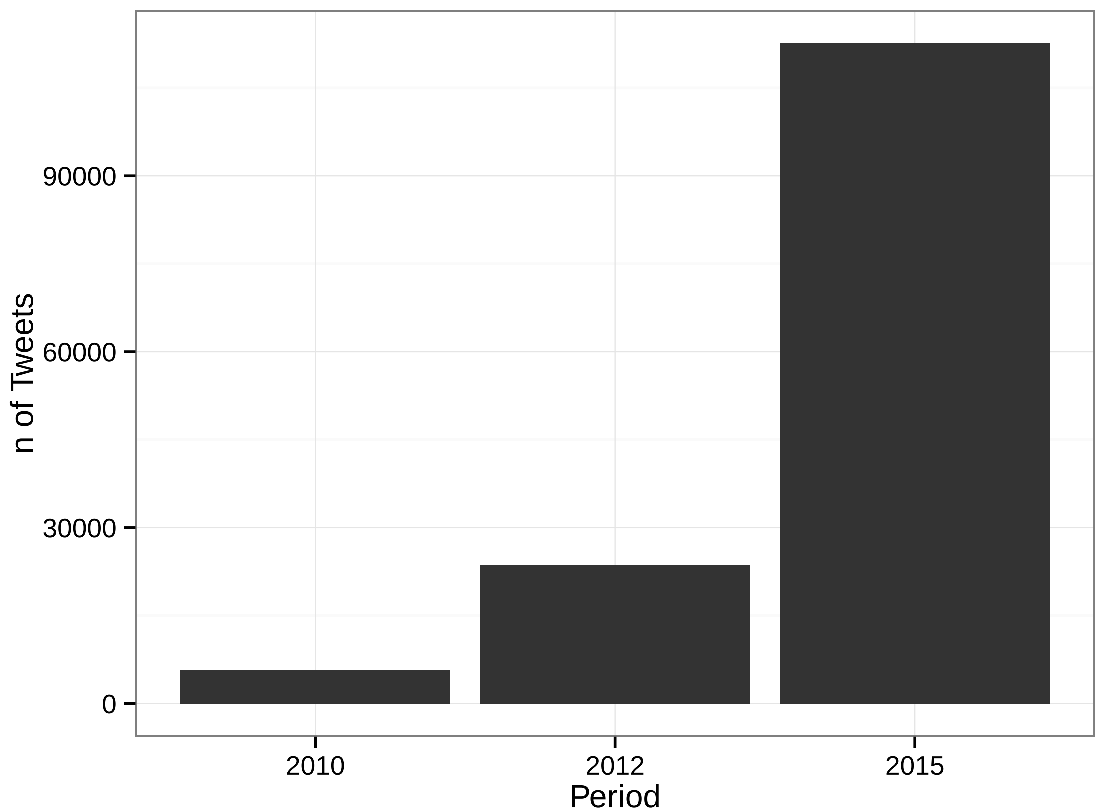

Figure 5: Tweets matching our queries by year

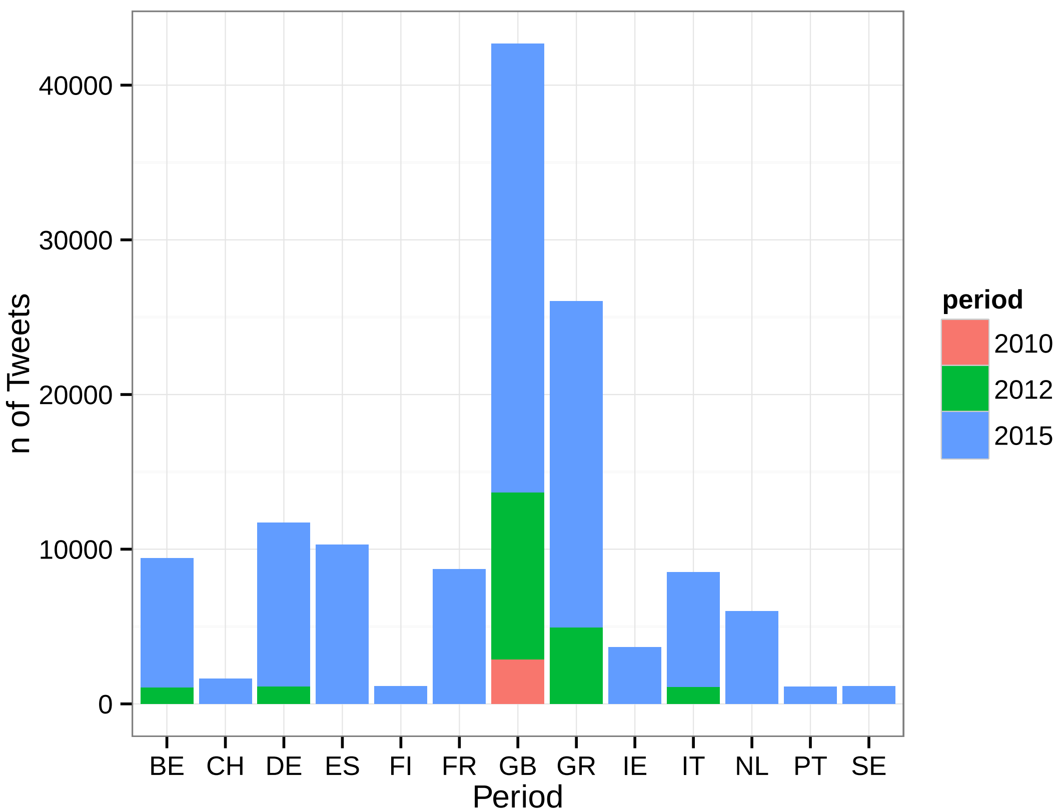

Figure 6: Spatio-Temporal distribution of total tweets in corpus

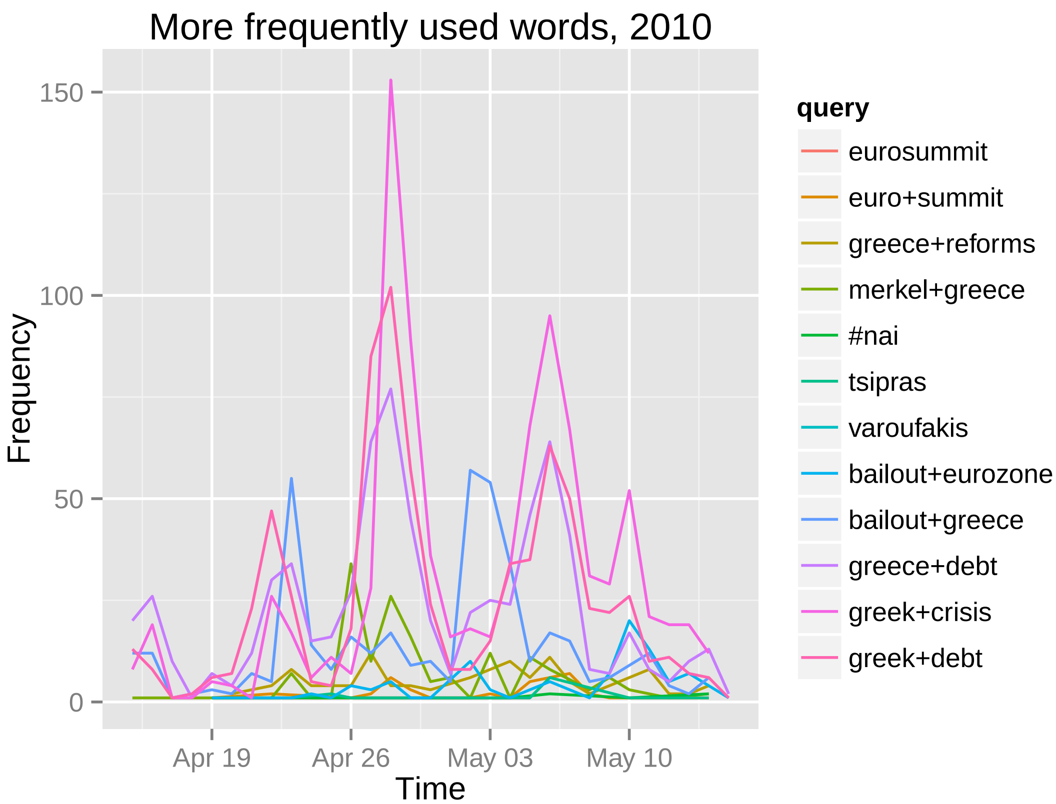

Figure 7: Most common search queries, 2010

on April 28th. However, the use of "greek+crisis" in tweets dropped dramatically for a short period of around 6 days, before it peaked again, however somewhat less pronounced. The drop in usage of "greek+crisis" was accompanied by an increase in usage of "greek+bailout", which peaked around May 2nd before it subsided. On May 2nd, the agreement on the First bailout package was reached (Goncalves 2015). During this short phase, "greek+bailout" was the most prominently used word combination (from our sample). Only after "greeck+crisis" became most frequently used words.

In the 2012 period, there was a discourse on the greek crisis weeks before the agreement. "greek+crisis" was used from the beginning of February until February 21st when the agreement between the Greek government and the Troika was reached (Economist 2015) constantly in between around 40 to 75 tweets. The usage of "greek+crisis" peaked at the date when the agreement was reached (in contrast to 2010 when it dropped around this phase). However, just like in 2010, "bailout+greece" became the most frequently used phrase in all tweets (of our sample) at the time when agreement was reached. After that, twitter users tweeted less about "bailout+greece" and "greek+crisis". However, it took nearly a month for those words being used only ten times a day. Notably, "grexit" which had not been used in 2010 was used frequently in the negotiation phase.

In the 2015 period, "greek crisis" was still amongst the most commonly used words in the corpus of tweets we collected. While at its highest peaks in 2010 and 2012 it was used around 120 times daily, we found around 1500 occurances of "greek+crisis" in one day in the 2015 period. Equally popular were "#ThisIsACoup" and "Varoufakis", and "Tsipras". The most freqquently used word in the 2015 period was "eurosummit" which was used a lot in the time running up to the summit and during the summit. During the summit, "greek+crisis" also peaked. In the aftermath of the announcement that an agreement had been reached, there is much less tweets. The Hashtag "This is a Coup" emerged on July 14th, peaked in July 15th - equally prominent as "greek+crisis". "Varoufakis" becomes frequently used in tweets in late July, which may have been caused by the release of an interview interview with revelations about a previously secret "plan B". (Kitsantonis 2015) "bailout+greece" peaked around the 14th of August, when the Greek parliament approved the third bailout package. Towards the

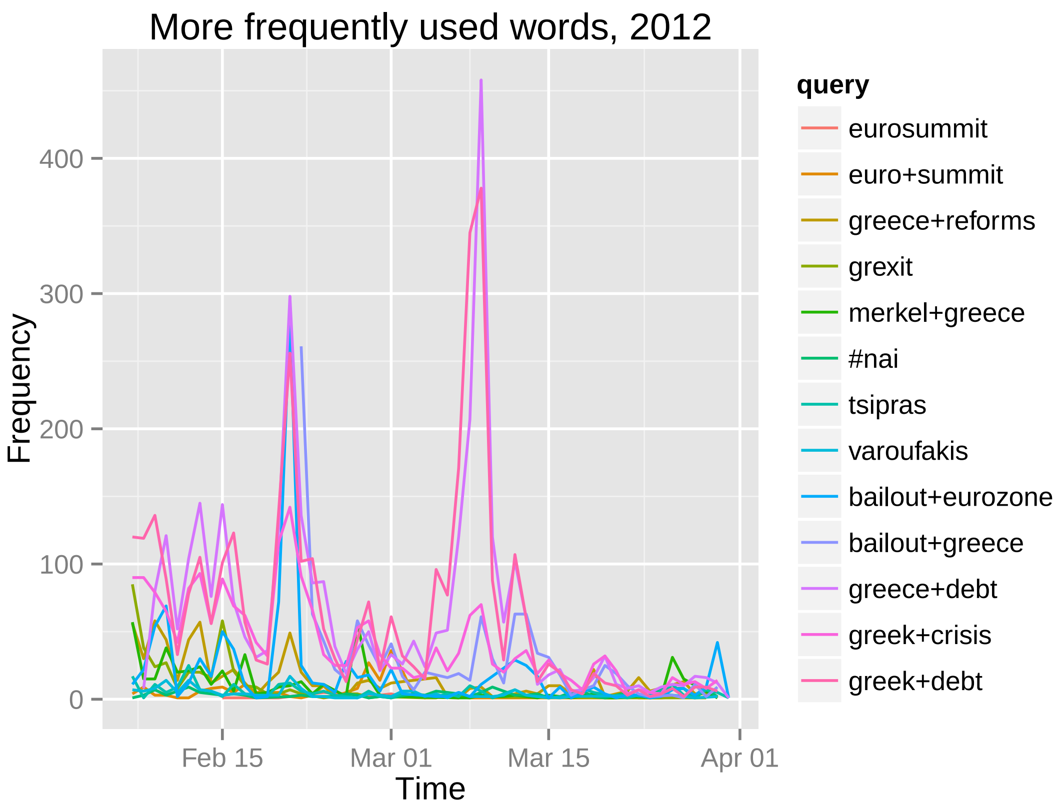

Figure 8: Most common search queries, 2010

end of the period, "varoufakis" and "tsipras" become frequently used. "varoufakis" has a second peak when he resigns (on August 20th).

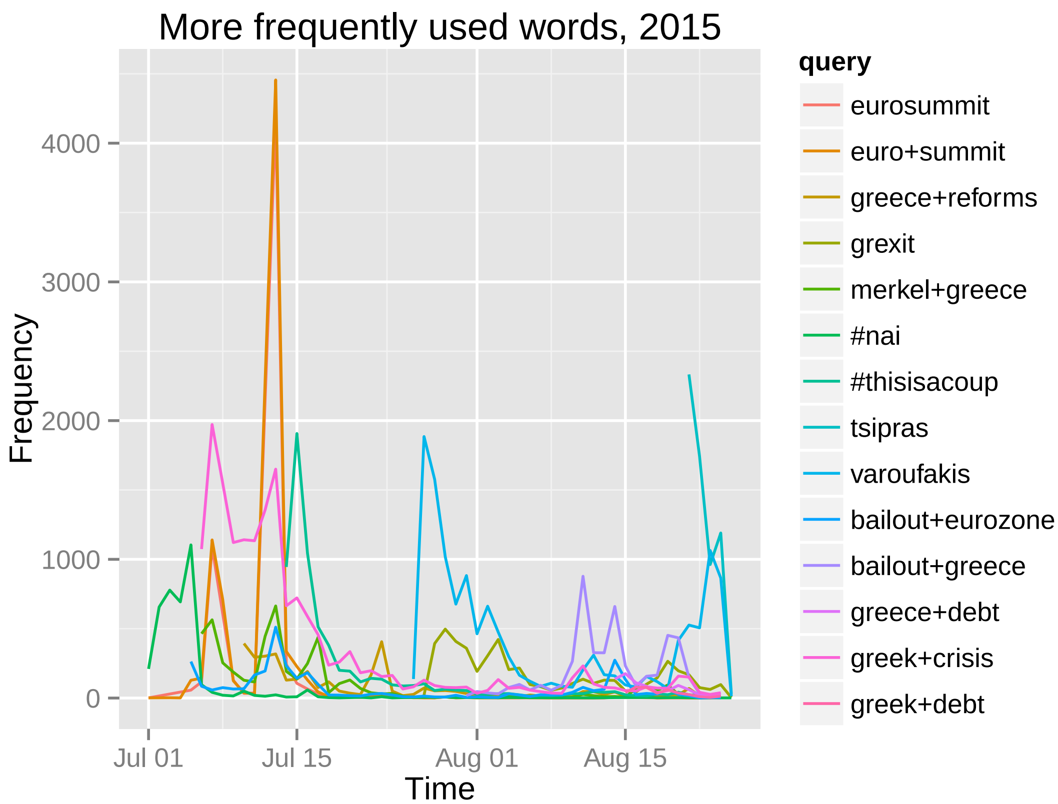

Figure 9: Most common search queries, 2010

Inferential Statistics: Sentiment Analysis

We summarised sentiment over time and across countries. Figures 10-12 show total sentiment scores (across all queries) and sentiment variance for all of the European countries in our dataset in each of the three time periods. We can visually identify variation in the overall score and in the variance between countries and across the three time periods.

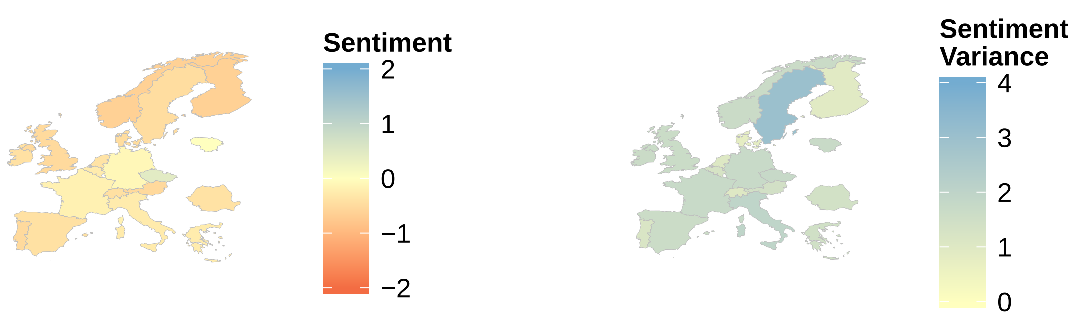

Figure 10: Sentiment scores and variance across all search queries (2010)

To test whether the spatio-temporal distribution of sentiment in tweets is significant, we regress sentiment score on country and time dummies.

An F-test on the joint significance of country and time finds that we can reject at the 1% significance level ($p = 0$) the null hypothesis that time-period and country have no effect on sentiment.

We focus our analysis in this paper on the whole corpus, though we also release our dataset aggregated at the country level in interactive form, where users can investigate specific queries during any of the periods.

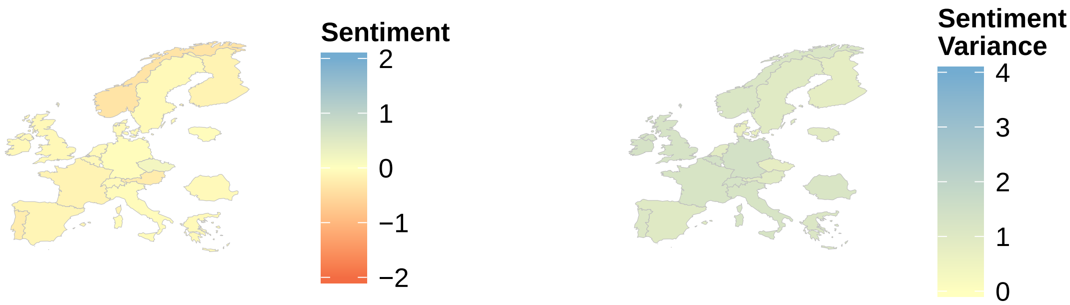

Figure 11: Sentiment scores and variance across all search queries (2012)

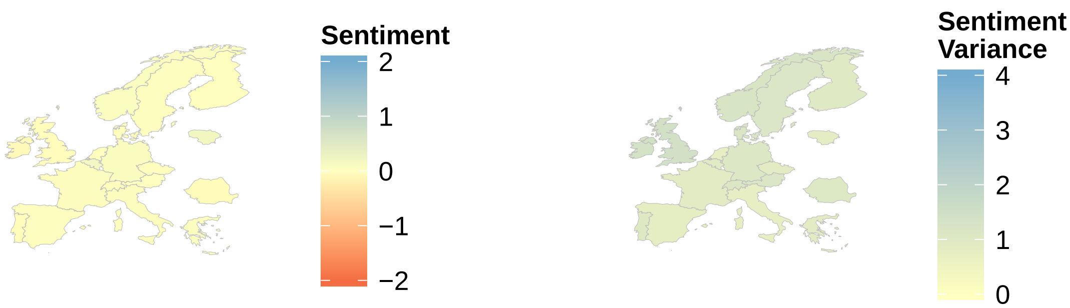

Figure 12: Sentiment scores and variance across all search queries (2015)

To focus our analysis, we zoom in to a single time period (2015), and to two countries where we expect sentiment to diverge: Greece and Germany. Figure 13 plots shows daily sentiment averages, surrounded by a band the width of the square root of the variance. A first observation is that the paths of the lines are remarkably similar. Following on from this, we can see how sentiment in both countries reacts to actual political events. On the 13th of July, Greece and its creditors struck a deal to agree on terms for bailout funds in exchange for stringent cuts and reforms. (Haenska and Bauchowitz 2015) identify the emergence of the "#thisisacoup" hashtag on the 13th. We can observe a clear drop in sentiment over the next two days in both countries. On the 14th of August the Greek parliament approved the package of reforms that formed part of the bailout deal. This time we see a spike in sentiment, once again in both countries, but more pronounced in Germany. Indeed, where each countries lines diverge, it is most often Greece which displays less positive sentiment

We once again regress sentiment on country and time, this time limiting out results to 2015 period and the two countries here.

Despite this, the countries' lines do sometimes diverge. This is particularly apparent here. This shows a difference in the average sentiment between the two countries at the same point in time. To establish that this difference is not due to random noise, we once again regress total sentiment score on country and time dummies, this time limiting the analysis to the two countries in question and to the selected query.

We looked at individual tweets and compared the results to our perception of sentiment of the tweet. While there seemed to be some problems, the sentiment analysis scores did seem reasonable.

However, when calculating the sentiments for queries, the returned results did not match our expectations. Below we report the resulted for the query "merkel+greece". We report the sum of the positive scores divided by the number of tweets. The figures indicate that the sentiments were very positive in most of the European countries and very similar.

We conclude from this that our current sentiment analysis does not return reasonable results. We will in the next stage of the project try to create a model that returns more accurate results.

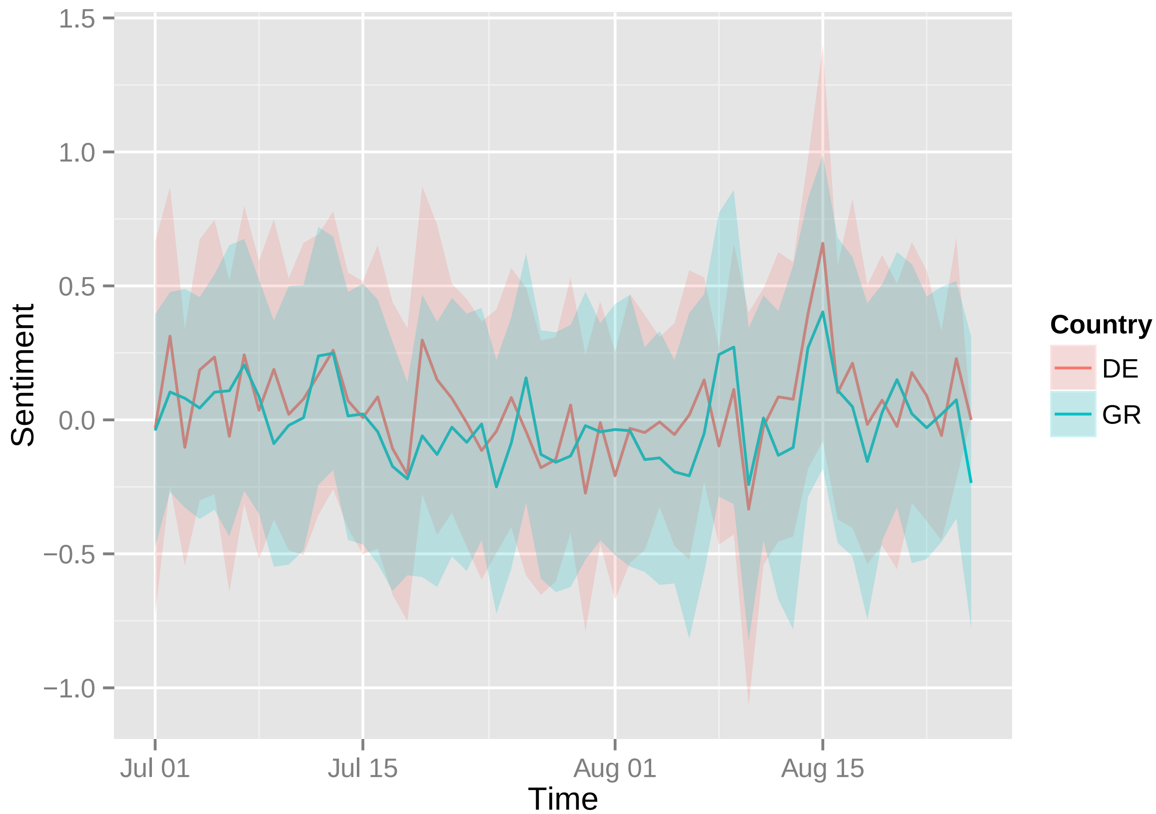

Figure 13: Sentiment scores (line) and variance (band) for all queries in 2015

Table 2: The country effect on twitter sentiment (controlling for day) in Greece and Germany in 2015

	Dependent variable:
	Tweet sentiment score
Greece	-0.038^{***}
	(0.012)
Constant (Germany)	-0.010
	(0.064)
Observations	31,893
R^2	0.020
Adjusted R^2	0.019
Residual Std. Error	0.983 (df = 31836)
F Statistic	11.808^{***} (df = 56; 31836)
Note:	*p<0.1; **p<0.05; ***p<0.01

Caveats

- Location
- translation: danish, romanian,
- "normalizaion" of values not conducted for some, because of huge data amounts
- SA: caution, however, about the utility of SA in predictive models (Grimmer and Stewart 2013, 4). "For shorter texts, accompanying information (or an extremely large volume of texts) is often necessary for classifcation or scaling methods to perform reliably" (Grimmer and Stewart 2013, 6)
- research design: case selection ambigious - for research on a common european sphere, choosing a case that by definition devides along national lines

Conclusions

References

Annau, Mario. 2014. *Tm.plugin.sentiment: Text Corpus Sentiment Analysis*. http://R-Forge.R-project.org/projects/sentiment/.

Balahur, Alexandra, and Marco Turchi. 2012. "Multilingual Sentiment Analysis Using Machine Translation?" *Proceedings of the 3rd Workshop in Computational Approaches to Subjectivity and Sentiment Analysis P. 52-60*. http://publications.jrc.ec.europa.eu/repository/handle/JRC73293.

Biernacki, Patrick, and Dan Waldorf. 1981. "Snowball Sampling: Problems and Techniques of Chain Referral Sampling." *Sociological Methods & Research* 10 (2). SAGE Publications: 141–63.

Drewski, Daniel, Jürgen Gerhards, and others. 2015. "Has There Been a European Public Discourse on the Euro Crisis?"

Economist, The. 2015. "Greece's Return to the Markets." Accessed December 3. http://www.economist.com/news/finance-and-economics/21600727-bond-issue-milestone-there-still-long-way-go-prodigal-son.

Gentry, Jeff. 2015. *TwitteR: R Based Twitter Client*. http://CRAN.R-project.org/package=twitteR.

Goncalves, Sergio. 2015. "Markets Calm Amid New Austerity for Portugal." Accessed December 3. http://www.reuters.com/article/2010/05/14/us-eurozone-idUSTRE6400PJ20100514.

Graham, Mark, Scott a. Hale, and Gaffney,Devin. 2014. "Where in the world are you ? Geolocation and language identification in Twitter." *The Professional Geographer* 00 (2013): 00. doi:10.1080/00330124.2014.907699.

Grimmer, Justin, and Brandon M Stewart. 2013. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts." *Political Analysis*. SPM-PMSAPSA, mps028.

Haenska, Max, and Stefan Bauchowitz. 2015. "A European Twitter Sphere? What Tweets on the Greek Bailout Say About How Europeans Interact Online." Accessed October 23. http://perma.cc/6ZVZ-X5GT.

Henrique, Jefferson. 2015. "Get Old Tweets Programmatically." Accessed October 21. https://github.com/Jefferson-Henrique/GetOldTweets.

Hornik, Kurt, Johannes Rauch, Christian Buchta, and Ingo Feinerer. 2015. *Textcat: N-Gram Based Text Categorization*. http://CRAN.R-project.org/package=textcat.

Hughes, Amanda Lee, and Leysia Palen. 2009. "Twitter adoption and use in mass convergence and emergency events." *International Journal of Emergency Management* 6 (May): 248. doi:10.1504/IJEM.2009.031564.

Jungherr, Andreas, Pascal Jürgens, and Harald Schoen. 2012. "Why the Pirate Party Won the German Election of 2009 or the Trouble with Predictions: A Response to Tumasjan, a., Sprenger, tO, Sander, PG, & Welpe, IM 'Predicting Elections with Twitter: What 140 Characters Reveal About Political Sentiment'." *Social Science Computer Review* 30 (2). SAGE Publications: 229–34.

Kitsantonis, Jack Ewing; Niki. 2015. "Greece Made Preparations to Exit Euro." Accessed December 3. http://www.nytimes.com/2015/07/28/business/greece-debt-varoufakis-recording.html?_r=1.

Lucas, Christopher, and Dustin Tingley. 2014. *TranslateR: Bindings for the Google and Microsoft Translation APIs*. http://CRAN.R-project.org/package=translateR.

Metz, Cade. 2015. "Twitter Now Lets You Search for Any Tweet Ever Sent." Accessed October 21. http://www.wired.com/2014/11/twitter-now-lets-search-tweet-ever-sent.

Narr, Sascha, Michael Hulfenhaus, and Sahin Albayrak. 2012. "Language-Independent Twitter Sentiment Analysis." *Knowledge Discovery and Machine Learning (KDML), LWA*, 12–14.

Pak, Alexander, and Patrick Paroubek. 2010. "Twitter as a Corpus for Sentiment Analysis and Opinion Mining." *In Proceedings of the Seventh Conference on International Language Resources and Evaluation*, 1320–26. doi:10.1371/journal.pone.0026624.

Risse, Thomas. 2003. "An Emerging European Public Sphere? Theoretical Clarifications and Empirical Indicators."

———. 2014. "No Demos? Identities and Public Spheres in the Euro Crisis." *JCMS: Journal of Common Market Studies* 52 (6): 1207–15. doi:10.1111/jcms.12189.

Scola, Nancy. 2015. "Library of Congress' Twitter Archive Is a Huge #FAIL." Accessed October 23. http://www.politico.com/story/2015/07/library-of-congress-twitter-archive-119698.html.

"The Search API." 2015. https://dev.twitter.com/rest/public/search. Accessed October 21.

Theussl, Stefan, Paul Hofmarcher, and Kurt Hornik. 2015. *Tm.lexicon.GeneralInquirer: General Inquirer Categories*.

Tornes, Adam. 2015. "Instant and Complete Access to Every Historical Public Tweet." https://blog.twitter.com/2015/full-archive-search-api.

Trenz, Hans-Jörg. 2015. "Europeanising the Public Sphere – Meaning, Mechanisms, Effects." In *Interdisziplinäre Europastudien*, edited by Ulrike Liebert and Janna Wolff, 233–51. Springer Fachmedien Wiesbaden. doi:10.1007/978-3-658-03620-1_11.

Tumasjan, Andranik, Timm Oliver Sprenger, Philipp G Sandner, and Isabell M Welpe. 2010. "Predicting Elections with Twitter: What 140 Characters Reveal About Political Sentiment." *ICWSM* 10: 178–85.

Vilares, David, Miguel A Alonso, and Carlos Gómez-Rodriguez. 2015. "Sentiment Analysis on Monolingual, Multilingual and Code-Switching Twitter Corpora." In *6TH WORKSHOP oN COMPUTATIONAL APPROACHES tO SUBJECTIVITY, SENTIMENT aND SOCIAL MEDIA ANALYSIS WASSA 2015*, 2.

Weller, Katrin, Axel Bruns, Jean Burgess, Merja Mahrt, and Cornelius Puschmann. 2013. *Twitter and Society*. Vol. 89. Peter Lang New York.

Zimmer, Michael. 2010. "Is It Ethical to Harvest Public Twitter Accounts Without Consent?" http://www.michaelzimmer.org/2010/02/12/is-it-ethical-to-harvest-public-twitter-accounts-without-consent/.

Zimmer, Michael, and Nicholas John Proferes. 2014. "A topology of Twitter research: disciplines, methods, and ethics." *Aslib Journal of Information Management* 66 (3): 250–61. doi:10.1108/AJIM-09-2013-0083.