

Identifying Networks in Social Media: The case of #Grexit

Georgios Magkonis¹ · Karen Jackson²

Published online: 29 March 2018

© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract We examine the intensity of ‘#Grexit’ usage in Twitter during a period of economic and financial turbulence. Using a frequency-analysis technique, we illustrate that we can extract detailed information from social media data. This allows us to map the networks of interest as it is reflected in Twitter. Our findings identify high-interest in Grexit from Twitter users in key peripheral countries, core Eurozone members as well as core EU member states outside the Eurozone. Overall, our study presents a useful tool for identifying clusters. This is part of a new research agenda utilising the information extracted from big data available via social media channels.

Keywords Networks · Big data · Twitter · Geo-location data · Grexit

1 Introduction

The role of social media has become progressively more important as a tool to spread information, including the analysis of current economic and business issues. Almost all large enterprises, government bodies and economic agents (financial commentators, major economists, etc.) are connected with at least one social media network. The usefulness of analysing social media data is well understood by financial experts (Tett 2013) as well as big corporations specialised on internet services, such as ‘google analytics’ and ‘yahoo finance’. Furthermore, data from social media like Twitter, Facebook, LinkedIn (Zhan et al. 2014; Fire et al. 2016; Moya-Gómez et al. 2017) as well as from other IT services (Westland et al. 2016) has recently become the focus of considerable research.

✉ Georgios Magkonis
georgios.magkonis@port.ac.uk

¹ Subject Group of Economics & Finance, Portsmouth Business School, University of Portsmouth, Portland Street, Richmond Building, Portsmouth PO1 3DE, UK

² Department of Economics & Quantitative Methods, Westminster Business School, University of Westminster, 35 Marylebone Road, London NW1 5LS, UK

The potentially important effect of online activity on economic and political outcomes has recently attracted the interest of academics. Specifically, Da et al. (2011) and Joseph et al. (2011) focus on the relationship between online search activity and US stock price movements, where their evidence suggests that price movements can be predicted on the basis of online research. Furthermore, Dergiades et al. (2014) consider the relationship between sovereign spreads and social media activity within the context of the Greek debt crisis. They find that there is bi-directional causality between Greek spreads and social media discussion. Apart from economic and financial research, the employment of online data has also been used in the political science sphere. Ko et al. (2014) explore Twitter dynamics during elections, while Berger and Morgan (2015) analyse the demographic profiles of ISIS-terrorist supporters on Twitter. At the same time, researchers recognise the need to introduce new rigorous methods to analyse big data (Monroe et al. 2015; Wang et al. 2017).

Social media data flows can be viewed as a network of interconnected nodes. The analysis of such networks may identify sub-structures/clusters, analogous to a core-periphery structure (Borgatti and Everett 2000). The connections between clusters within social media networks suggest that the structure cannot be strictly divided into separate groups (Pattison 1993). The methodology adopted in this paper does not impose further restrictions on how nodes are linked to one another. Our approach is a novel addition to the umbrella of techniques referred to as network analysis and responds to the call for new methods to analyse big data.

We study a specific piece of information and how it appeared in social media, more precisely on Twitter. The buzz word that assists us to extract messages from social media is 'Grexit'. This term is a relatively new word in financial vocabulary and it refers to the possibility of Greece leaving the Euro area. Grexit made its appearance during the European debt crisis. The Greek counterpart of the crisis began in late 2009 when the budget deficit estimates were revised to a value (12.5%) higher than the originally expected. This revision immediately threatened Greek fiscal credibility and resulted in its access to capital markets becoming significantly more difficult over time. In April 2010, it became evident that Greece was no longer able to finance its debt. One month later the Greek government and the so-called Troika (the European Commission, the IMF and the ECB) agreed the first bailout programme of €110 billion. In February 2012, the Greek government signed a second bailout package of €130 billion. The increasing threat of Grexit was reflected in public discourse as well as social media. This trend culminated in the general elections of May and June 2012. After a short period of political calm, Grexit reemerged as a significant possibility during the general elections of January 2015, where the new government doubted the conditions of the bailout agreements. Furthermore, the expectation of Grexit became higher as the term began to be used officially by European authorities.

Economic theory has emphasised the role of expectations and how these can lead to self-fulfilling prophecies via network effects (Krugman 1996; Chang and Velasco 2001 and Burnside et al. 2004). This research was mainly motivated by the need to understand the currency crises of 1990s and early 2000s. Given the fact that such prophecies have played a role in significant economic outcomes, we propose a method of measuring the intensity of public expectations of future events using information from social media. Specifically, we empirically investigate the geography of the 'Grexit'-term usage in Twitter. We focus on the Twitter network, due to its voluminous

amount of data that is rapidly updated (Laney 2012). Furthermore, Twitter is used for brief official announcements by government bodies, private institutions and individuals. Identifying the geography of networks in several settings has recently become a popular topic (see for instance, Illenberger et al. 2013, Holl and Mariotti 2017).

In this study, we use as a clustering algorithm a frequency analysis technique. More precisely, we use Winger function analysis to identify the range and uniformity of locations involved in Twitter discussions regarding Grexit. In doing so, this paper contributes to the methodological toolkit of big data network analysis. Application of this technique to cross-sectional data indicates a broader usefulness of time-frequency techniques, where the Wigner function is part of this analytical grouping. Our results are in accordance with the economic developments of each country. The remaining paper is structured as follows; Section 2 describes the methodology; Section 3 presents the data collection process; Section 4 presents the results and Section 5 concludes.

2 Methodology

Time-frequency analysis is an established technique used across different disciplines among social sciences (Aguiar-Conraria et al. 2012; Rua and Nunes 2009; Caraiani 2012). More recently this analysis has been deployed within finance and economics, where the technique has typically been used to explore cyclical data using Wavelet tools (Aguiar-Conraria and Soares 2014). The Wavelet function integrates to zero, suggesting movements above and below the x -axis (i.e. cyclical movements). However, time-frequency analysis can be applied to cross-sectional data using the Wigner function, which integrates to one (Earnshaw et al. 2012). Wigner function analysis uses a pseudo-probability function, which is particularly useful in identifying uniform or non-uniform trends. This tool quickly analyses and compresses information, which makes this technique useful as part of the methodological toolkit of big data network analysis.

This paper uses data collected from social media and replaces the ‘time’ variable with alternative variables/orderings.¹ This type of big data can be analysed using Winger Function analysis, such that we shed light on the neighbourhood/s of locations linked to tweeting about Grexit as well as the uniformity of tweets within neighbourhoods. Turning to briefly define the model, the real and positive function $f(x)$ is defined for integer values of x . With interpolation and the assumption that outside this finite interval $f(x)$ goes very fast to zero (or alternatively that $f(x)$ is a periodic function of x), it can be defined for all real values of x . Furthermore, we introduce the real function $F(x)$ as

$$F(x) = \left[\frac{f(x)}{A} \right]^{\frac{1}{2}} \quad (1)$$

where,

$$A = \int_{-\infty}^{\infty} dx f(x) \quad (2)$$

¹ The use of frequency analysis to explore ordered categorical variables is outlined in Earnshaw et al. (2012).

The Fourier transform of $F(x)$ is given by

$$\tilde{F}(\nu_x) = \int F(x) \exp(i\nu_x x) \quad (3)$$

The Wigner function $W(x, \nu_x)$ is defined as

$$W(x, \nu_x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dx' F\left(x - \frac{x'}{2}\right) F^*\left(x + \frac{x'}{2}\right) \exp\left[i\left(x' \nu_x\right)\right] \quad (4)$$

$F^*(x)$ is the complex conjugate of $F(x)$. ν_x is the frequency corresponding to the variable x , and plays an important role in our methodology. It simultaneously provides information for the function $F(x)$ and its Fourier transform $\tilde{F}(\nu_x)$. The areas of the x - ν_x plane where $W(x, \nu_x)$ is large signal strong activity/interest. The clusters of different levels of interest constitute the mapping of social media networks.

3 Data and Research Design

The twitter data used in this analysis was collected over the period 16-28 February 2015. These dates constitute a period where there was an active European discourse debating the possibility of Grexit. On the 20th of February there was a preliminary agreement between the new Greek government and the Troika and, subsequently, the possibility of Grexit was significantly reduced.² The collected tweets were identified by the buzzword/hashtag #Grexit, where the collection process made use of the Twitter Search API. To facilitate the data handling process we used the TAGS (Twitter Archiving Google Sheet) tool. This uses the scripting capability of Google Spreadsheets to regularly collect any tweets containing a given search term (<https://tags.hawksey.info/>).

The second step consists of data extraction and cleaning. The aim of this process is, firstly, to narrow the dataset in order to include only EU countries and, secondly, to identify from where the tweets originate. Ideally, we would use independently verified data for location. However, the origin is imperfectly recorded since the user location is subject to user-input bias. This bias is two-fold. Firstly, a user may input a location that does not represent where she/he is resident. Secondly, the location field is not populated from a set of pre-defined set of alternatives. Instead, the user manually types their perception of their own location (or in many cases it is left blank).³

Therefore, we use two proxies for user location: location (as inputted by the user) and user language (selected by the user from a pre-defined list). These two proxies provide us with two separate datasets. The first proxy reduces the number of tweets in the dataset since a significant proportion of users do not specify a location. The total

² While we were writing this study, a number of significant events took place (e.g., financial turmoil in Greek economy in June and July 2015). These events made Grexit even more probable. We decided to retain the analysis for February agreement for two reasons. Firstly, we had already started working on the paper. Secondly, if we were to consider the period of turmoil, which lasted for at least two months, the amount of data was unhelpfully excessive. This would make the data handling almost impossible.

³ For example, a twitter user may use 'free-text' to input their location as London, England, UK or the United Kingdom.

tweets recorded during the period were 64,004 and from these 7858 indicated a location that was reported and could be accurately filtered as originating from a particular EU country (Table 1). However, this is not a significant problem since 7858 is a large number of observations. The second proxy, user language, requires matching to a location based on the country where the language is most commonly used (Table 2). Hence, through this process we make the data identifiable (Shlomo and Goldstein 2015).

In the previous section, we mentioned that our use of cross-sectional data requires the identification of a non-time ordering. For robustness, this paper will adopt two alternative orderings. Both cases provide a measure of economic connectedness to the Greek economy. We select variables that summarise trade and financial linkages. Trade is measured as Greek imports from EU countries. The left column of Table 3 shows the

Table 1 Tweets based on location

User-defined location	Frequency of tweets
UK	1417
Austria	127
Belgium	217
Bulgaria	24
Croatia	12
Cyprus	403
Czech	26
Denmark	46
Estonia	19
Finland	50
France	548
Germany	817
Hungary	2
Italy	319
Latvia	5
Lithuania	5
Luxembourg	27
Malta	19
Netherlands	351
Poland	39
Portugal	152
Romania	24
Slovakia	7
Slovenia	21
Spain	180
Sweden	104

The number of tweets by user-defined location is based on the cleaned dataset, where multiple user-inputted locations have been assigned to a particular country. The Republic of Ireland has been excluded due to difficulties in accurately filtering the data. Data from Greece are excluded from the analysis since the target of the paper is to analyse non-Greek interest.

Table 2 Tweets based on language

Language - explanation	Frequency of tweets	Country (association)
Catalan; Valencian	114	Spain
Czech	42	Czech Republic
Danish	129	Denmark
German	7342	Germany
English	36,707	
Spanish; Castilian	3110	Spain
Basque	6	Spain
Finnish	92	Finland
French	2882	France
Galician	21	Spain
Hungarian	5	Hungary
Italian	2954	Italy
Dutch; Flemish	3214	Netherlands
Polish	114	Poland
Portuguese	363	Portugal
Romanian; Moldavian; Moldovan	5	Romania
Swedish	189	Sweden
Total tweets	57,289	

6715 out of 64,004 tweets were excluded due to unclear, Greek or non-EU language selection. English has been dropped from the sample due to multiple country associations. The removal of English reduces any potential bias, as English tweets can come from many countries

Greece's trade partner starting from the most to least important. The second ordering variable is the Credit Default Swap (CDS). This variable can be viewed as a proxy of financial fragility for each EU economy. A high (low) CDS value reflects an increasing (decreasing) cost of borrowing from financial markets. We assume that a financial fragile economy would be more seriously affected if Greece were to leave the Eurozone. In the right column of Table 3, we report the ordering according to CDS.

We describe the data with a real and positive function $f(x)$ of a variable x which takes values $x = 1, \dots, N$. In this paper, x describes the location of the information where each integer value of x represents a country and $f(x)$ represents the number of tweets originating from country x quoting #Grexist. The dataset based on final language, contains the variable x that takes on the values $1, \dots, 13$ representing EU countries: Czech Republic, Denmark, Finland, France, Germany, Hungary, Italy, Netherlands, Poland, Portugal, Romania, Spain and Sweden. In a similar vein, the dataset based on location takes on the values $1, \dots, 25$ and $1, \dots, 26$ for the ordering of Greek imports and CDS, respectively.

4 Results and Discussion

Our analysis provides four sets of results, which use combinations of the two proxies for user location and the two ordering variables. The results of our analysis using 3-D

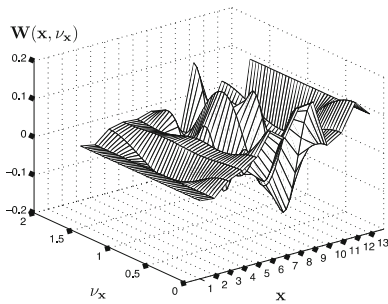
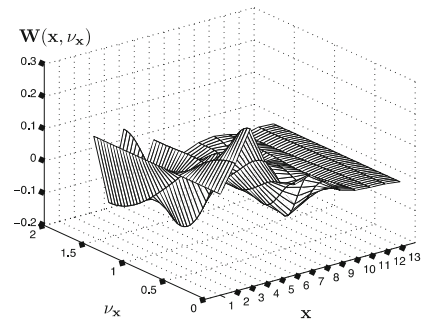
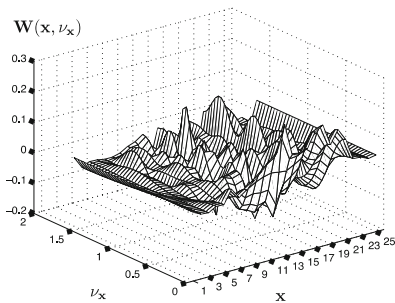
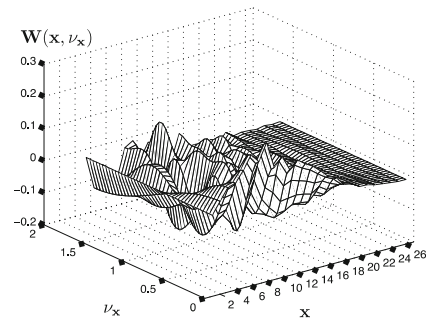
Table 3 Ordering Variables

Greek imports from...	Credit Default Swap (CDS)
Germany	Cyprus
Italy	Croatia
Netherlands	Bulgaria
France	Portugal
Spain	Hungary
Bulgaria	Slovenia
Belgium	Romania
United Kingdom	Italy
Romania	Lithuania
Poland	Latvia
Austria	Spain
Denmark	Poland
Cyprus	Slovak Republic
Hungary	Estonia
Sweden	Czech Republic
Czech Republic	France
Slovenia	Belgium
Slovak Republic	UK
Portugal	Luxembourg
Finland	Denmark
Croatia	Finland
Luxembourg	Netherlands
Malta	Austria
Lithuania	Sweden
Latvia	Germany
Estonia	

Greek import data (measured in thousands of US dollars) is sourced from UNComtrade. CDS data are collected from Datastream. For both cases, countries are listed from the largest value to smallest. In the case of CDS, Malta is excluded due to data availability

plots are presented in Fig. 1, where panels A) - D) show alternative datasets. The full set of results, including contour maps and results matrices are available upon request. The two sets of results based on language (Fig. 1 panels A) and B)) cover a narrower range of countries than those based on location (Fig. 1 panels C) and D)). Therefore, we have four sets of results for the common set of 13 countries.

Clusters of high values of the Wigner function allow us to identify six countries (out of 13) where Twitter users exhibit what we will refer to as ‘high interest’ in Grexit: Denmark, France, Italy, Netherland, Poland and Spain. There are two further countries (out of 13) that are on the ‘periphery’ (i.e. outer range) of the high value Wigner function clusters: Germany and Romania. Henceforth we will refer to these two countries as having Twitter users exhibiting ‘medium interest’. Thirdly, we have a further three countries where users exhibit high interest but they only appear in two

(a) Final language, CDS**(b)** Final language, Greek imports**(c)** Location, CDS**(d)** Location, Greek imports**Fig. 1** $W(x, \nu_x)$ by x variable, ordering variable

datasets based on the location variable: Austria, Belgium and UK. They will also be referred to as ‘high interest’. Finally, there are countries where there is no data or limited interest exhibited by twitter users. The results are summarised in Fig. 2.

A cursory glance at the basic statistics (for example, number of tweets per country) as compared to our results, provides a clear illustration that Wigner function analysis extracts additional information from the data that is otherwise likely to remain hidden. For example, the number of tweets from Germany as well as their rank in the ordering variables can be contrasted to the results from the Wigner function analysis. Our analysis provides consistent country grouping that are insensitive to the chosen ordering, Greek imports or CDS. This is an important robustness check since our research departs from the approach of using a time ordering, commonly used in the literature. Furthermore, our categorisation of countries, and the twitter users linked to the locations, reflects the nature of activity within the time-period. In other words, the twitter users contained in our dataset actively communicated via twitter rather than passively reading updates. In addition, the period in question followed sustained media coverage. Therefore, the users actively engaging with Twitter to communicate are likely

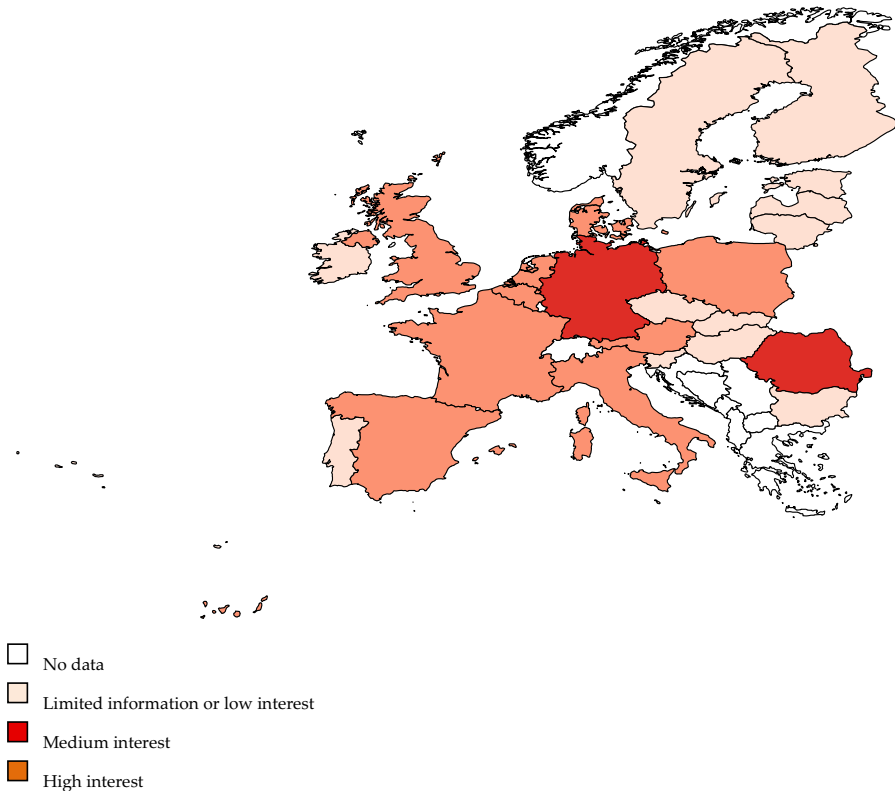


Fig. 2 Mapping of European Clusters

to have been exhibiting behaviour consistent with significant interest in Grexit. It is feasible that this is derived from self-interest as the likelihood of Grexit may have more significant implications for some members of the EU rather than others.

This is reflected in our results. Firstly, Italy and Spain are the largest economies of the European periphery. However, these countries are considered as economically weak and unlikely to withstand further shocks. Even though neither economies have participated in a bail-out programme they are under the tight supervision of the European authorities. Both financial analysts as well as the general public believe that the financial turmoil from Grexit would quickly spread to Italy and Spain (Micossi 2015). Similarly, France has also been required to adopt careful macroprudential policies in order to respond to concerns regarding its performance against the Maastricht criteria and the Stability and Growth Pact (Bennani et al. 2014).

Furthermore, we find an intensive interest originating from Belgium, Netherlands, Austria as well as three countries that they have not adopted Euro; UK, Denmark and Poland. Starting from the British interest, this result is in accordance with the on-going discussions regarding the referendum concerning exit from the EU. Our evidence suggests that the British public is closely following developments regarding Grexit, as the outcome of the referendum may be influenced by the events in Greece. Greek departure from the monetary union could significantly influence the final decision of British voters.

On the other hand, the finding regarding Denmark can be explained by the participation of the Danish Krone in the Exchange Rate Mechanism II (ERM II). This requires the Danish central bank to keep the exchange rate of Danish Krone to the Euro free-floating within the narrow band of $\pm 2.25\%$. Given Denmark's trade dependence on the Eurozone, volatility in the exchange rate created by Grexit could have significant economic consequences. Moreover, Poland has particular significance since it is the second largest non-Eurozone EU economy. Prior to the threat of Grexit, the majority of the Polish public was consistently in favour of adopting Euro (European Commission 2014). However, discussions of the possibility of Grexit shifted Polish public opinion towards a more Eurosceptic stance. This was a major influencer during the Polish general elections during late 2015 (Foy 2015).

Furthermore, Austria, Belgium and Netherlands are classified as countries that belong to the core of the Eurozone. In this sense, the public interest regarding a serious event in the periphery, which may lead to a change in its structure, is also expected to be significant. In particular, as far as Belgium is concerned, its public debt is estimated around 106.6% (European Commission 2015a). This has raised considerable discussion regarding debt sustainability and fears of the need of a bail-out. In addition, these concerns have been raised in the European Commission's Fiscal Sustainability Report (European Commission 2016).

The countries that have Twitter users exhibiting intense interest can be understood in terms of economic fundamentals. Turning to Germany, there is a key difference. Germany has held a central political role in discussions regarding Grexit. However, there is less evidence that Germans should be concerned about a significant destabilising impact from Grexit. Given that the German economy is the 'warehouse of Europe', the German sovereign state has benefited from a gradually declining cost of borrowing from the financial markets. This provides the confidence that the German economy is robust to a variety of financial shocks and hence our result of medium interest. Finally, tweets from Romania also show a moderate interest. According to the 2015 Eurobarometer (European Commission 2015b), the Romanian public are strongly in favour of Euro adoption. This suggests that Romanians view completing EU membership, by joining the monetary union, as less economically worthwhile despite the turmoil in Greece. Nevertheless, Romanians remain concerned about the prospect of Grexit due to the number of Greek subsidiary banks operating in the country (Wheatley 2015).

Overall, results obtained by Wigner function analysis show that there are two groups of countries. As shown above, the outcomes are robust across alternative orderings. They are driven by the likely transmission of financial contagion from Grexit as well as by the economic developments in each country.

5 Conclusions

Social scientists have begun to realise that the information posted on social media can be useful for understanding social and economic phenomena. This paper has applied a method that can be used to extract further information regarding the public perception of key economic/policy changes. In this paper, we examine the Web traffic of Twitter users from different locations across Europe. We find that interest is most intense when

there is a more significant risk of a negative economic impact from Grexit. This case study also illustrates that this technique has a broader usefulness as part of the group of techniques referred to as network analysis. Furthermore, policy makers can use the findings from this type of modelling as an early warning system for shifting public opinion. In terms of future research, this framework can be applied to explore clusters of interest regarding Brexit.

References

- Aguiar-Conraria L, Soares MJ (2014) The continuous wavelet transform: moving beyond uni- and bivariate analysis. *J Econ Surv* 28:344–375
- Aguiar-Conraria L, Magalhães PC, Soares MJ (2012) Cycles in politics: wavelet analysis of political time series. *Am J Polit Sci* 56:500–518
- Bennani T, Després M, Dujardin M, Duprey T, Kelber A (2014) Macroprudential framework: key questions applied to the French case. Occasional papers, 1
- Berger J, Morgan J (2015) The ISIS twitter census: defining and describing the population of ISIS supporters on twitter. Brookings Project US Relat Islamic World 3:20
- Borgatti SP, Everett MG (2000) Models of core/periphery structures. *Soc Networks* 21:375–395
- Burnside C, Eichenbaum M, Rebelo S (2004) Government guarantees and self-fulfilling speculative attacks. *J Econ Theory* 119:31–63
- Caraiani P (2012) Money and output: new evidence based on wavelet coherence. *Econ Lett* 116:547–550
- Chang R, Velasco A (2001) A model of financial crises in emerging markets. *Q J Econ* 116:489–517
- Da Z, Engelberg J, Gao P (2011) In search of attention. *J Financ* 66:1461–1499
- Dergiades T, Milas C, Panagiotidis T (2014) Tweets, google trends, and sovereign spreads in the GIIPS. *Oxf Econ Pap* 46:1–27
- Earnshaw RA, Lei C, Li J, Migassabi S, Vourdas A (2012) Large-scale data analysis using the wigner function. *Physica A* 391:2401–2407
- European Commission (2014) Directorate-general economic and financial affairs. Flash Eurobarometer 400, Introduction of the Euro in the more recently acceded member states
- European Commission (2015a) European economy, macroeconomic imbalances country report-Belgium 2015. Occasional Paper 212
- European Commission (2015b) Directorate-general economic and financial affairs. Flash Eurobarometer 418, Introduction of the Euro in the more recently acceded member states
- European Commission (2016) European economy, fiscal sustainability report 2015. Institutional Paper 018
- Fire M, Puzis R, Elovici Y (2016) Organization mining using online social networks. *Netw Spat Econ* 16:545–578
- Foy H (2015) Candidates put euro at centre of polish presidential race. *Financial Times*, 1 April
- Holl A, Mariotti I (2017) The geography of logistics firm location: the role of accessibility. *Netw Spat Econ* 18:1–25
- Illenberger J, Nagel K, Flotterod G (2013) The role of spatial interaction in social networks. *Netw Spat Econ* 13:255–282
- Joseph K, Wintoki MB, Zhang Z (2011) Forecasting abnormal stock returns and trading volume using investor sentiment: evidence from online search. *Int J Forecast* 27:1116–1127
- Ko J, Kwon HW, Kim HS, Lee K, Choi MY (2014) Model for twitter dynamics: public attention and time series of tweeting. *Physica A* 404:142–149
- Krugman P (1996) Are currency crises self-fulfilling? NBER Macroeconomics Annual 1996, Volume 11. MIT press, Cambridge
- Laney D (2012) The importance of ‘big data’: a definition. Gartner, Stanford
- Micossi S (2015) What future for the eurozone? <http://www.voxeu.org/article/future-eurozone-0>
- Monroe BL, Pan J, Roberts ME, Sen M, Sinclair B (2015) No! Formal theory, causal inference, and big data are not contradictory trends in political science. *PS Polit Sci Polit* 48:71–74
- Moya-Gómez B, Salas-Olmedo MH, García-Palomares JC, Gutiérrez J (2017) Dynamic accessibility using big data: the role of the changing conditions of network congestion and destination attractiveness. *Netw Spat Econ* pp 1–18. <https://doi.org/10.1007/s11067-017-9348-z>
- Pattison P (1993) Algebraic models for social networks. Cambridge University Press, Cambridge

- Rua A, Nunes LC (2009) International comovement of stock market returns: a wavelet analysis. *J Empir Financ* 16:632–639
- Shlomo N, Goldstein H (2015) Editorial: big data in social research. *J R Stat Soc Ser A* 178:787–790
- Tett G (2013) Markets insight: wake up to the Twitter effect on markets. *Financial Times*, 18 April
- Wang XB, Cao X, Yin K, Adams TM (2017) Modeling vehicle miles traveled on local roads using classification roadway spatial structure. *Netw Spat Econ* 17(3):713–735
- Westland CJ, Hao JX, Xiao X, Shan S (2016) Substitutes, complements and network effects in instant messaging services. *Netw Spat Econ* 16:525–543
- Wheatley J (2015) Spectre of Grexit sparks fears for central and eastern Europe. *Financial Times*, 21 June
- Zhan X, Ukkusuri SV, Zhu F (2014) Inferring urban land use using large-scale social media check-in data. *Netw Spat Econ* 14:647–667