

Chicken or Egg? The reciprocal influence of press and politics

Jan Kleinnijhuis
The Network Institute
VU University Amsterdam
The Netherlands
j.kleinnijhuis@vu.nl

Wouter van Atteveldt
The Network Institute
VU University Amsterdam
The Netherlands
wouter@vanatteveldt.com

Antske Fokkens
The Network Institute
VU University Amsterdam
The Netherlands
a.s.fokkens@vu.nl

Abstract

Do the media cover the considerations and actions of Congress with respect to the financial crisis, or do politicians follow the lead of the media? This paper shows how relevant information about the functioning of press and politics can be acquired using keyword search, VAR analysis and granger causality. Media lead Congress, but no influence is found for problem diagnosis and solutions. We conclude with a discussion on how more advanced NLP techniques can be used to answer remaining questions on political use of the media.

1 Introduction

Media portray what is going on in politics, but politics also responds to what is being discussed in the media. We investigate the overlap and the temporal ordering of political discussions in the US Congress and in US and European media.

- What is the relation/overlap between topic coverage, associations and perceptions presented in political documents and that presented in the media?
- Does increased attention for a topic in politics precede an increase in the media and/or vice versa (Granger causality)?

This paper aims not to present a methodologically novel way of analyzing textual material. Rather, it aims to show how simple keyword analysis can be combined with established statistical methods and a theoretical understanding of the topic to yield interesting insights about the financial crisis. Using keyword analysis, we identify topics of discussion (political and economic issues) and actors (institutions, representatives) involved. Subsequently, their associations

(as inferred from their co-occurrence) and evaluations (as inferred from a financial sentiments analysis) are identified. Finally, a time series analysis using Vector Auto Regression (VAR) and granger causality is conducted to infer whether topics in politics precede the media or vice versa.

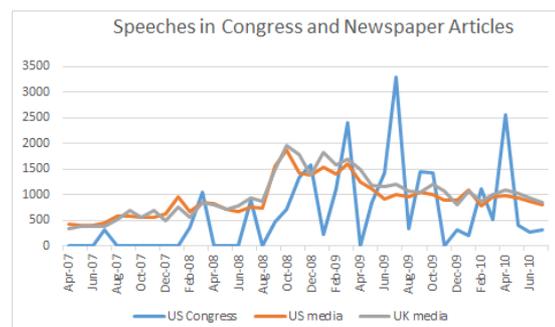
2 Data and method

This paper analyzes written material from two sources: US congressional hearings (n=24,989 speeches/interventions) and US and UK media. The former were provided as part of the competition, while the latter were retrieved from LexisNexis and Factiva based on a query of 64 terms related to the financial crisis and 85 terms related to the financial sector. Documents were selected from April 2007 until July 2010 (from a few months before the financial crisis until the signing of the Dodd–Frank Act).

For the US media (n=36.958 articles in total), we included the NY Times (n=16.904), Washington Post (n=11,740) and USA Today (n=3,802). UK media (n=39.453) consisted of The Times (n=17419), The Guardian (n=13787), and The Mirror (n=8247, inc. Sunday Mirror).

Figure 1 shows the distribution of articles/speeches over time in the three venues. Immediately apparent is that congressional attention depends on the congressional agenda, with in

Figure 1: Overview of analyzed documents



many months no meeting at all. Newspaper attention develops more smoothly. US newspapers and UK newspapers co-vary over time to a large degree. Newspaper attention peaks in the month after the collapse of Lehman Brothers (15 September 2008). Congressional attention peaks in the month after president Obama launched the first version of a Wall Street reform act.

2.1 Methods

In earlier research into the media coverage of the financial crisis a hierarchical ontology was developed based on a categorization of actors (e.g. US actors, subdivided in the White House, Congress, the FED, etc., EU actors, etc.), financial actors (e.g. USA banks, including investment banks, UK banks, etc.) and on a categorization of topics. Topics were divided in problems (e.g. housing crisis, financial crisis, euro crisis), causes (e.g. speculation, bonuses), consequences (e.g. bankrupt banks, bankrupt nations, unemployment), solutions (e.g. austerity politics (e.g. budget cuts), regulation, including the Dodd Frank Wall Street Reform Act).

Latent Dirichlet was applied to retrieve clusters of concepts from (lemmas in) congressional hearings and newspaper data. The results showed that the structure of the top-down ontology, but enabled a refinement of several concepts, e.g. a more refined classification of FED committees and of terms involved in the housing crisis.

To tap financial sentiments we started from search queries for concepts applied in earlier research (e.g. hope/fear, gain/loss, up/down; Kleinnijenhuis et al. 2013) and from Stone et al.'s (1965) General Inquirer word lists that were applied in financial research by Teflock (2007) and further adapted by Loughran and McDonald (2011). Van Dalen (2014) showed that the uncertainty index of these authors is an even better predictor in financial analysis than their general positive/negative indices.

The final ontology consists of 7 actor categories, with 41 subcategories, 4 topics (problems, causes, consequences, solutions; subdivided further into 28 issues) and evaluations (subdivided further in 21 evaluative dimensions).

Search queries were used to retrieve hits for 102 concepts (41+28+21 subcategories, plus searches for their unions to retrieve hits for the 7+4+1 main categories. Attention for (sub)topics was measured with its number of hits. Asymmetric associations of a with b were measured as their conditional probability (cf. Tversky, 1977; Ruigrok and Van Atteveldt, 2007).

3 Congress-Media: attention overlap?

Table 1 gives an overview of the percentage of attention for actor categories. US Congress concentrates on the USA, as do the US media. The amount of attention for US actors in UK media exceeds the amount of attention for EU actors in US media. Although the London City was the largest center of financial trade at the start of the crisis, attention for UK actors in the US is relatively low. UK newspapers pay special attention to the UK, but also to European countries that ran into financial troubles (the PIIIGS nations) especially to nearby Ireland. In UK newspapers attention for the FED is larger than for the ECB. UK media also pay a lot of attention to the precise role of banks.

According to a seminal article by Entman (1993), political discourse is framed in terms of specific problems-solution combinations. Table 2 shows the distribution of attention for various issues, headed under the topics problems, causes, consequences and solutions. The foremost important conclusion from Table 2 is that US Congress concentrates on the *causes* and *solutions*, whereas the media remain focused on the crisis *problems*. Whereas newspapers talk relatively often about banks (cf. Table 1), US Congress addresses the *causes* of failed banks, especially speculation (22.0%, e.g. trade in derivatives and securities, credit default swaps, short selling). Bonuses that may have caused risky behavior of greedy bankers are a major issue both in Congress and in US and UK newspapers (22.8 .24.1%). Congress concentrates on new financial regulation as the *solution* for the crisis. Budget cuts in government expenditures are more often discussed in UK newspapers.

These conclusions appear to be robust for slightly different categorizations. It should be noted, for example, that a categorization of issues as either causes or consequences is too rigid (debts in 2009 and 2010, for example, were largely a consequence, rather than a cause of the crisis). The most important evaluative dimension is uncertainty.

Equal percentages in Tables 1 and 2 for Congress, US media and UK media may still result from attention to the same issues at very different points in time. Correlation coefficients per month (n=40 months) for each actor/issue separately may tell us whether attention of Congress, US media and UK media is synchronous (correlation coefficient near +1) or a-synchronous (near -1).

Table 1: Attention for actors (% of total hits)

	US Congress	US media	UK media
USA	80.9	72.8	23.7
FED	19.9	7.1	2.4
EU	0.4	1.7	5.4
ECB	0.2	0.8	1.4
EUCountries	0.5	4.1	31.3
UK	0.3	0.9	24.4
France	0.1	1.3	2.8
Germany	0.1	1.2	2.2
PIIGS	0.2	2.7	10.1
Iceland	0.0	0.3	1.3
Greece	0.0	0.7	1.3
Ireland	0.1	0.4	3.9
Portugal	0.0	0.2	0.3
Spain	0.0	0.5	1.6
Italy	0.0	0.5	1.6
Banks	10.2	13.1	21.6
Other FinancialActc	3.5	1.8	1.9
Investors	4.2	3.8	6.0
N of hits all actors	72698	309246	204446

Overall these correlation coefficients show synchronous attention. Tables 1 and 2 reveal considerable disparities between Congress on the one hand and the media on the other, but the data reveal correlation coefficients >0.5 (=synchronization) for financial actors, issues such as bonuses and recession, and evaluations/emotions such as anger, blame and hope. US media and UK media reveal very high correlations (=synchronous attention) for almost every actor and almost every issue.

4 Congress-Media: who follows whom?

Normative political theory maintains that in a representative parliamentary democracy media attention should follow attention by elected parliamentarians. Vector Auto Regression (VAR) is used to ascertain whether attention in Congress precedes or follows attention in the media. We applied a pooled VAR-analysis (for $n=102$ pools of separate actors/issues) to monthly data (April 2007-July 2010; $n=40$ months). The irregular agenda of congressional meetings implies that a shorter time span to exert influence (e.g. days, weeks) is dubious. Figure 2 shows cross-lagged influences that denote Granger-causality as arrows between Congress, US media and UK media. Autoregression coefficients are plotted as reflexive arrows. Arrows are labeled with regression coefficients that indicate the direction and the strength of influence. Boxes are labeled with numbers that denote the proportion of explained

Table 2: Attention for actors (% of total hits)

	US Congress	US media	UK media
Problem	20.5	38.4	41.4
housingCrisis	4.4	3.2	1.6
Financial \$ crisis	9.9	19.8	16.0
EuroCrisis	1.3	4.6	10.6
recession	0.8	4.6	6.1
Cause	51.3	36.2	36.6
Speculation	22.0	6.3	4.6
Bonus	22.8	23.2	24.1
Debt	5.4	6.5	7.4
Consequence	7.5	12.9	10.4
BankruptBank	0.0	0.2	0.3
BankruptState	4.2	4.5	2.2
PublicMood	1.5	4.2	2.7
Unemployment	0.8	1.9	2.2
Inflation	0.8	1.2	1.8
Solution	20.7	12.5	11.6
Regulation	20.0	9.9	6.9
BudgetCuts	0.1	0.9	2.8
N of hits all topics	119983	458982	414401

variance by (cross-lagged and autoregressive) incoming arrows.

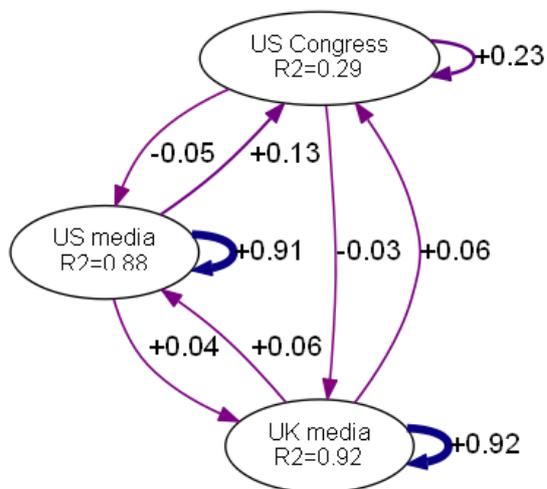
The most important conclusion from the pooled Vector Auto Regression is that US Congress is influenced more by the media, even more by US media (+0.13) and by foreign media (+0.06) than the other way around: regression coefficient are slightly negative but still significant.

An important side remark is that the congressional agenda is more innovative than the agendas of US media and UK media, which can be seen from the relatively low proportion of explained variance ($R^2=0.29$ for Congress as compared to $R^2=0.88$, resp. 0.92 for the media). We do not know how this picture would have been different, if the Presidential agenda and the FED agenda would have been included in the model. Figure 2 shows further that the agendas of the US media and UK media are determined largely by their own history. For the US Congress the autoregression coefficient is less strong (+0.23) because of agenda pressure due to which financial issues will not be on the congressional agenda in a lot of months (cf. Figure 1).

4.1 Who follows whom for specific topics

To answer our research question we need to know whether the overall pattern of influence as depicted in Figure 2 holds also at the deeper level of associations of actors and issues towards

Figure 2: Who follows whom?
Results of the pooled VAR Analysis



specific other issues and evaluations. The time series for associations with each specific topic and evaluation is used as input for detailed VARs to estimate whether Congress follows the media or vice versa.

For most topics and evaluations a weak bidirectional pattern of influence coefficients results, which implies that associations with most topics and most evaluative dimensions do not have a clear origin. Although Congress pays more attention to causes and to solutions than the media (cf. Table 1), it is not the case that Congress directs the attention of the media towards associations with causes or consequences one month after a month with excessive congressional attention. The reason could be that the causes of the financial crisis (the nature of debts and speculative trading in high frequency markets) and their solutions (precise prescriptions for financial institutions) are supposed to be too complex and too technical for media consumers. This indicates that US politics is unable to raise substantial attention for problem solution starting from a causal diagnosis.

However, for problems we see a clear causal pattern: associations in the media of specific actors/issues with problems precede these associations in Congress to a stronger degree (+0.20) than the other way around (+0.06). The same causal pattern holds for associations with the credit crisis (+0.17 as compared to +0.07). Associations with the evaluative dimension of blaming show again the same causal pattern (+0.20 as compared to +0.08). By and large these results are in line with the pessimistic view that problem awareness and putting blaming on others traverse from the media to politics.

Unexpectedly, a pattern with a stronger influence of associations in Congress on associations in the media (+0.16) than the other way around (+0.09) emerges for the evaluative dimension of Hope. Although the media do not let themselves be led by Congress when highly technical political solutions are at stake, hope in Congress is a light on the path of the media. Hopes and fears in the media appear to be important in shaping the economic future (cf. Kleinnijenhuis et al., 2008).

5 Discussion

Keyword analysis and time series models were used to establish to what degree politics or media are leading in the discourse about the financial crisis. Whereas the media are occupied by problems, Congress devotes attention to their causes and their possible solutions. The overall influence of the media on Congress exceeds the reverse influence. The media exert a stronger influence on associations with problems, especially with the credit crisis and with blaming Congress, than the other way around. No clear pattern is discerned when it comes to the strongholds of Congress: the causal diagnosis of problems and their subsequent solution. Associations with hope in Congress do exert a relatively strong influence on hopes according to the media.

This shows that using relatively simple methods substantively interesting information can be extracted from large sets of digitally available text. What is missing here is the political dimension: which politicians introduce or comment on specific diagnoses or solutions? And do these politicians discuss these matters primarily in Congress, do they use the media as a tool for communicating with their constituency and/or colleagues, or are they used by the media to provide quotes and official provenance for the stories the journalists want to tell? To answer these questions, we need to segment the text corpora into the texts uttered by individual politicians or parties. Moreover, especially in a complex topic such as the crisis we need to distinguish between whether something is cause or effect of another variable (e.g. debts can be both cause and result of the crisis, and the diagnosis and solutions will differ radically based on this role). This requires more in-depth analysis than provided by keyword analysis, and we will investigate this angle in the near future using a rule based approach on top of syntactic analysis (Van Atteveldt 2014a; 2014b).

References

- De Nooy, Wouter, & Kleinnijenhuis, Jan. (2013). Polarization in the media during an election campaign: a dynamic network model predicting support and attack among political actors. *Political Communication*, 30(1), 117-138. doi: DOI: 10.1080/10584609.2012.737417
- Dixon, Robert M.W. (1992). *A new approach to English grammar, based on semantic principles*. Oxford: Clarendon.
- Dixon, Robert M.W. (2005). *A semantic approach to English grammar*. Oxford University Press.
- Kleinnijenhuis, Jan, & de Nooy, Wouter. (2013). Adjustment of issue positions based on network strategies in an election campaign: A two-mode network autoregression model with cross-nested random effects. *Social Networks*, 35(2), 168–177. doi: 10.1016/j.socnet.2011.03.002
- Kleinnijenhuis, Jan, de Ridder, Jan A., & Rietberg, Ewald.M. (2008). Reasoning in economic discourse: an application of the network approach to the Dutch press. In K. Krippendorff & M. A. Bock (Eds.), *The Content Analysis Reader* (pp. 430-442). Thousand Oaks: Sage.
- Kleinnijenhuis, Jan, Schultz, Friederike, Oegema, Dirk, & van Atteveldt, Wouter. (2013). Financial news and market panics in the age of high-frequency sentiment trading algorithms. *Journalism*, 14(2), 271-291. doi: 10.1177/1464884912468375
- Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*, 66(1), 35-65.
- Ruigrok, Nel, & van Atteveldt, Wouter. (2007). Global Angling with a Local Angle: How U.S., British, and Dutch newspapers Frame Global and Local Terrorist Attacks. *The Harvard International Journal of Press and Politics*, 12, 68-90. doi: 10.1177/1081180X06297436
- Stone, P. J., Dunphy, D. C., Smith, M. S., & Ogilvie, D. M. (1966). *The General Inquirer: A Computer Approach to Content Analysis*. Boston: The MIT Press.
- Tetlock, Paul C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, 62(3), 1139-1168.
- Tversky, Amos. (1977). Features of similarity. *Psychological Review*, 48(4), 327-352.
- Van Atteveldt, Wouter, Kleinnijenhuis, Jan, & Ruigrok, Nel. (2008). Parsing, Semantic Networks, and Political Authority: Using Syntactic Analysis to Extract Semantic Relations from Dutch Newspaper Articles. *Political Analysis*, 16(4), 428-446. doi: 10.1093/pan/mpn006
- Van Atteveldt, Wouter, Sheaffer, Tamir, and Shenhav, Shaul (2014a), Semantic Network Analysis of Frame Building during war: Mediated Public Diplomacy in Gaza, Georgia, and Iraq, presented at ISA 2014, Toronto
- Van Atteveldt (2014b), Quotes as Data: Extracting Political Statements from Dutch Newspapers by applying Transformation Rules to Syntax Graphs, presented at MPSA 2014, Chicago
- Van Dalen, Arjen. (2014). The impact of economic news in uncertain times. A time-series analysis of economic coverage in Denmark 1996-2012. *Paper prested at ASCOR research lab, 14 january 2014*.