RESEARCH

Open Access

Mapping discursive dynamics of the financial crisis: a structural perspective of concept roles in semantic networks

Adina Nerghes[1*†], Ju-Sung Lee[2†], Peter Groenewegen[1] and Iina Hellsten[1]

*Correspondence:
adina.nerghes@vu.nl
[†]Equal contributors
[1]Department of Organization
Sciences, VU University, Amsterdam,
De Boelelaan 1081, 1081 HV
Amsterdam, The Netherlands
Full list of author information is
available at the end of the article

Abstract

Background/purpose: Convenient access to vast and untapped collections of documents generated by organizations is a highly valuable resource for research. These documents (e.g., press releases) are a window into organizational strategies, communication patterns, and organizational behavior. However, the analysis of large document corpora requires appropriate automated methods for text mining and analysis that are able to take into account the redundant and predictable nature of formalized discourse.

Methods: We use a combination of semantic network analysis and network centrality measures to overcome these particular challenges and to explore the dynamic structural space of concepts in formalized documents pertaining to the recent financial crisis.

Data: For our analyses, we collect the press releases of the European Central Bank (ECB) and the United States' Federal Reserve System (Fed) issued between 2006 and 2013 in order to examine their semantic networks before, during, and after the recent financial crisis. Their press releases are notably impactful in their influence on other financial institutions and society at large, especially during times of financial volatility.

Results: The structural space created from joint centrality metrics reveals salient shifts in the discursive practices of the ECB and Fed. In particular, the Fed exhibits greater attentiveness to the financial crisis especially during the crisis itself, while the ECB's attention is delayed and increasing steadily. Furthermore, we show both the Fed's and the ECB's discourse transitioning into a new "hybrid" state, rather than returning to the pre-crisis status quo.

Conclusions: Examining the semantic networks of organizational text documents, we find that our analytic approach reveals important discursive shifts, which would not have been discovered under traditional text-analytic approaches. We demonstrate the utility of this approach in investigating large text corpora of organizational discourse, and we anticipate our methods to be comparably valuable in the analysis of a large spectrum of formal and informal discourse.

Keywords: Text; Semantic networks; Centrality; Discourse; Structural space; Financial crisis

Nerghes *et al. Computational Social Networks* (2015) 2:16

Page 2 of 29

Introduction

The increasing availability of online textual information opens new venues for large-scale research into organizational discourse and vocabulary shifts of organizations [1]. In particular, numerous text documents regarding organizational activities and objectives are generated daily across the world. However, large corpora of such text documents are difficult to analyze without proper methods which are in part automated.

Another research challenge is that organizational documents are often manifestations of highly formalized discourse, which leads to redundant, structured, and even predictable [2] language use. Discourse "acts as a powerful ordering force in [and by] organizations" [3] because meaning is negotiated in organizations, and these meanings shape organizational practices [4], as well as the external presentation of organizations. As a carrier of power, language—commonly labeled as discourse—has the ability to order and constitute the social world [3]. By content and structure, discourse signals consequential information to other organizations and society in general. Its timely analysis may be crucial in order to understand the dynamic character of such signals, yet this analysis is often challenging. The approach employed in this study has been designed to deal with complex semantic networks generated from large text corpora of formal organizational discourse (i.e., press releases). More precisely, the method assesses dynamic discursive shifts in complex semantic networks, highlighting the crucial distinction between connective and popular concepts.

Different approaches to text analysis include information retrieval, lexical analysis to study word frequency distributions, pattern recognition, tagging/annotation, information extraction, and data mining techniques. Direct text-analytic methods such as word frequencies and term frequency-inverse document frequency (tf-idf) scores are limited in the subtlety of their inferences, revealing an incomplete picture of the discursive dynamics within the text. Without intending to minimize the utility and effectiveness of other text-analytical approaches, we propose that semantic network analysis becomes an invaluable and time-efficient tool for exposing subtle patterns of large text corpora. In this paper, we employ semantic network analysis to dissect organizational discourse in a structured manner. Specifically, we demonstrate how this approach reveals key aspects of the relations between words within as well as across documents and ultimately global shifts in organizational discourse above and beyond what direct text analysis would reveal.

Semantic network analysis is one of the areas of research that has gained popularity in recent years. This type of analysis maps networks of concepts (i.e., a concept being a word or multiple words) in the form of networks of meaning. Successful application of semantic network analysis to a variety of corpora addressing a multitude of contexts attests to the flexibility and adaptability of this method. A few examples can be found in the works of Leydesdorff and Welbers on co-words in context [5]; Doerfel and Barnett on the structure of the International Communication Association (ICA) [6]; Shim, Park, and Wilding on nuclear energy policy frames [7]; Danowski on mapping publics of a business [8]; Carley and Kaufer on symbols and symbolic activity [9]; Grebitus and Bruhn on the perception of food quality [10]; Kleinnijenhuis, Hoof, Oegema, and Ridder on the effects of news on voting [11]; Kim on Internet discourses of Korean supporters of Hwang Woo Suk [12]; Leydesdorff and Hellsten on the role of language in tracking the way stem cell research is represented in various contexts [13]; Leydesdorff and Hellsten on metaphors

and diaphors in scientific controversies [14]; and Nerghes, Hellsten, and Groenewegen on the evolution of metaphor families in media reports [15].

Although language can be suitably represented as a network of co-occurring words [16], semantic networks are often large and complex and exhibit highly intricate network structures [17–19]. Some posit these networks to exhibit stylized topologies such as *small-world* or *scale-free* [16, 18–21]. Semantic networks provide insights into how language serves as a framework for representing and communicating information. The complexity of large semantic networks arises not only from the size of the corpora but also from an array of global and local features, which in turn emerge from the structure of links between the concepts.

Secondly, our paper applies an approach for assessing dynamic shifts in formal discourse through the structural positions of semantic network nodes. The structural space method, introduced in earlier work [15, 22–24], combines two classic social network analysis structural measures degree (i.e., popularity) and betweenness centrality (i.e., connectivity) of concepts to create four structural roles for network nodes. The idea of structural roles in semantic networks is analogous to developments in social networks in which various approaches have been explored over the years. Examples are structural holes [25], equivalence [26–29], blockmodels [30, 31], and role structure [32] which have distinct functions in social theories of structure. However, the identification of structural roles through the combination of structural measures has not been widely explored in semantic networks yet. One such effort comes from Carley and Kaufer [9], which combines density, conductivity, and consensus to explore connectivity in semantic networks. A recent study [7] examines structural roles of concepts alternatively using raw centrality measures. The paper of Huang, Lv, Zhang, Yang, Zheng, and Wen [33] proposes a combination of three strongly correlated social network analysis (SNA) metrics (degree, ego-betweenness centrality, and eigenvector centrality) to evaluate only those top-ranked nodes (core nodes and bridge nodes) in undirected binary networks. The distinctive feature of our approach is the identification of four structural roles based on the combination of two structural measures, and thus, it is not merely focused on node popularity.

The corpora used in this study comprises the press releases issued by the European Central Bank (henceforth ECB) and the United States' Federal Reserve System (henceforth Fed) between 2006 and 2013. The ECB and the Fed determine the monetary policy for two of the world's largest currency areas. Established by the Treaty of Amsterdam in 1 June 1998 [34], the ECB is the formal successor of the European Monetary Institute. As one of the seven institutions of the European Union, the ECB is the central bank for the euro and administers the monetary policy of the 17 EU member states (using the euro), which constitute the Eurozone.

The ECB distributes large volumes of information (e.g., press releases, policy deliberations, public speeches, and annual reports) as one of their key policy tools. Because the ECB's only formal instrument, through which they can exert an (indirect) effect on asset prices (of key importance to the economy), is the overnight interest rate, their communications become a powerful tool. These can impact developments in the financial markets [35–38] and directly influence private sector expectations and are used to signal interest rate changes [39, 40]. The communications of the ECB also increase the predictability of interest rate decisions [41], being generally considered trustworthy and understandable by the public [42].

Nerghes *et al. Computational Social Networks* (2015) 2:16

Page 4 of 29

The Fed has a longer history than that of the ECB. Its establishment on December 23, 1913, was a response to a series of financial panics, particularly the panic of 1907. The duties of the modern Fed include conducting the nation's monetary policy, supervising and regulating banking institutions, maintaining the stability of the financial system, and providing financial services to depository institutions, the U.S. government, and foreign official institutions [43]. Among the central banks, the Fed has a unique structure due to a mix of private and public elements which serve the interests of the general public and the interests of private bankers.

Just as in the case of the ECB communications, the Fed's communications contain useful information about future monetary policy [44]. Studies have found that the Fed communications moved financial markets in the intended direction and that these communications are even more relevant to the markets during financial turmoil [45].

Although the two organizations serve similar roles in their respective economies, one study suggests that the ECB has moved its policy rate much less frequently than the Fed and that their interest rate behavior was rather different [46]. However, both organizations have had to face the effects of the recent financial crisis.

Many studies investigating central bank communications have at least three findings in common: (1) central bank communications regarding economic projections and monetary policy developments have significant impacts on the financial markets, (2) these effects are even greater when the communication channel is more formal, and (3) the more prominent the position of the communicator, the stronger the financial market reaction [38, 47, 48] (Reinhart, Vincent and Brian Sack. Grading the Federal Open Market Committee's Communications 2006. Unpublished). As such, the communications of central banks are of great importance in times of market volatility and financial crises. In this paper, we apply semi-automated text analysis (semantic network analysis) to unravel over time changes in formal discourses and thus contribute to insight in their flexibility.

The remainder of this paper is organized as follows. In the following section, we describe our data sets and the approach we are employing. The "Results" section presents the results of our analyses, and the "Conclusion" section summarizes our overall findings and discusses the benefits and limitations of our approach.

Data and methods

For this study, a total of 3010 press releases issued by the ECB (825 documents) and the Fed (2185 documents) between January 2006 and December 2013 have been collected from their web archives. For each organization, the collected press releases have been divided into four time periods each spanning a period of 2 years. The aggregation of data in these four periods was motivated by our aim of assessing the discursive shifts of the ECB and Fed at the different phases of the financial crisis. The first subsample covers the period just prior to the financial crisis: January 2006 until December 2007. We label this period *pre-crisis*. The second subsample (*crisis*) includes press releases issued between January 2008 and December 2009. The third subsample contains press releases issued between January 2010 and December 2011 and represents the *post-crisis* period. Lastly, the fourth subsample includes press releases issued between January 2012 and December 2013, further referred to as *recovery*. Table 1 presents a summary of the number of documents collected (*D*), the word counts (*W*), and the sentence counts (*S*) for each

Nerghes *et al. Computational Social Networks* (2015) 2:16

Page 5 of 29

Table 1 Press releases collected

	ECB			Fed		
	Documents	Words	Sentences	Documents	Words	Sentences
Pre-crisis	184	59,544	2133	386	80,844	5104
Crisis	203	66,349	2286	786	140,959	7464
Post-crisis	210	70,914	2528	635	104,762	4869
Recovery	228	88,323	3274	425	93,662	4236
$\sum$	825	285,130	10,221	2232	420,227	21,673
$\mu(W/D)$		343.88	12.32		193.53	10.09

organization and each period as well as summary statistics: sum of the counts (Σ) and the words and sentences per document ratio averaged across the periods (i.e., $\mu(W/D)$ and $\mu(S/D)$).

We note some differences between the corpora of the Fed and ECB. In particular, the Fed data contains a higher volume of text (i.e., total word count). However, the Fed's much higher document count translates to its press releases being smaller in word size and sentence count than those of the ECB; the difference in word counts per document between the two institutions is significant ($p < 0.001$ for a t-test and $p < 0.05$ for a Kolmogorov-Smirnov test). This points to a striking difference in the communication styles between the U.S. and European institutions and raises additional, interesting questions and hypotheses about implications of these statistics which we leave for future exploration.[1] While the higher volume of the Fed data is expected to confer larger semantic networks of prominent concepts, we expect the relationship to be more nuanced than a mere function of word count given the complex structure of language.

In this paper, we use semi-automated coding of concepts to be included in the semantic networks [8, 9, 49]. Each of the data samples *(pre-crisis, crisis, post-crisis,* and *recovery)* has been pre-processed with AutoMap [50]. The pre-processing removed all the noise words (e.g., numbers as words, verbs, and single letters) in the data and prepared it for the generation of semantic networks. A total of eight semantic networks (one for each organization for the four phases of the crisis) were generated using the software (see Table 2 for the descriptive statistics of each network). The generation of networks is based on Carley's approach to coding texts as cognitive maps [51] and Danowski's approach to proximity analysis [52].

Semantic networks translate text into networks of concepts and the links between them, in which a concept may be a word or a phrase (i.e., n-gram) [53]. The links between concepts are based on co-occurrence. For example, if two words co-occur within the specified window size, a link (or semantic network edge) will be formed. The window size determines the range of text words in which connections will be made between words within

Table 2 Descriptive statistics of each semantic network generated[a]

	ECB				Fed			
Measure	*Pre-*	*Cri-*	*Post-*	*Rec-*	*Pre-*	*Cri-*	*Post-*	*Rec-*
Node count (n)	580	627	648	755	827	1009	886	860
Sum of links	113,389	131,326	159,091	199,352	178,782	345,885	250,854	260,873
Density	0.205	0.201	0.216	0.191	0.147	0.155	0.152	0.150

Pre- pre-crisis, *Cri-* crisis, *Post-* post-crisis, *Rec-* recovery
[a] Each network is undirected, symmetric, and valued; only nodes with frequencies $\geq$10 have been included in the networks

the window [54]. The window size used for this study was one sentence. The value of strength for each link is determined by the frequency of co-occurrence [55].

As mentioned above, a concept in our semantic networks can be a single word or an n-gram. N-grams are coded by replacing the spaces between words with an underscore [50]. An example of such a conversion is "interest rate" being coded as "interest_rate". This procedure helps us identify the most common multi-word expressions used in text documents. Thus, when we refer to key concepts, we refer to single words as well as n-grams.

In Table 2, some summary statistics of the generated semantic networks are reported. The sum of links is the sum of all the values of the weighted edges/links, while the (unweighted) density expresses the proportion of non-zero (or non-null) edges to the count of possible edges in the undirected semantic network (i.e., $n/(n(n-1))$). These descriptive statistics show that even after employing a frequency threshold (≥ 10) the resulting networks are complex and relatively dense with high link counts.[2] Again, the statistics reveal some striking differences between the two institutions, this time in their semantic networks. The node counts and edge weight sums increase monotonically over the periods (i.e., time) for the ECB, while these exhibit distinct peaks at the *crisis* period for the Fed. This finding is congruent with the knowledge that the United States economy was both the source of and most affected by the financial crisis. That is, we expect more activity in the Fed press releases than in the ECB's during the *crisis* period. In fact, the ratio of the node and edge weights during the *crisis* against the averages across the periods is below 1.0 for the ECB and above 1.0 for the Fed.

The higher link weights yet lower, unweighted density of the Fed (compared to those of the ECB) appear contradictory. However, this finding indicates a higher repetition of concept co-occurrences but proportionally fewer unique dyads (i.e., pairs of linked concepts) in the Fed corpora. Whether this is due to differences in language use or policy aims between the institutions remains a research question worth exploring.

Interestingly, we find that the node counts are best predicted (statistically) by the sentence counts in Table 1 more so than the document or word counts which leads to a hypothesis about the structure of the language in our corpora: that unique concepts arise out of distinct sentences rather than other textual boundaries. The adjusted R^2 for nodes predicted separately by documents, words, and sentences are, respectively, 0.85, 0.87, and 0.91.

The combination of the complexity of these networks and the formal character of the documents from which they have been extracted poses a challenge for the analyst. To overcome this challenge, we propose using a structural space approach that considers total degree centrality and betweenness centrality of concepts in semantic networks, concurrently.

Centrality in networks

Even after decades of social network research, the current thinking about network centrality is still mostly defined by the work of Freeman [56] and Bonacich [57]. In 1977, Freeman developed a set of centrality measures based on betweenness [58]. In a follow-up article 2 years later, Freeman [56] elaborates on three concepts of centrality in a social network, which have since been further developed into degree centrality, closeness centrality, and betweenness centrality. The fourth commonly used measure, eigenvector centrality,

Nerghes *et al. Computational Social Networks* (2015) 2:16

Page 7 of 29

was popularized by Bonacich [57] and is based on previous graph-theoretical research [59, 60]. We now define and briefly elaborate on (total) degree centrality and betweenness centrality, the focal measures in this paper. Increasing interest in network centrality has been spurred by the growth of research on large-scale networks such as the hyperlinks of the worldwide web [61] and other online networks.

Total degree centrality

Total degree centrality is one of the most commonly used centrality measures in social network analysis [62]. The degree centrality of a node in a network reflects the number of other nodes incident to the focal node [56] (or, in the case of weighted networks, the sum of the weights of all the incident links) and thus measures the involvement of a node in its local network. Nodes with low total degree centrality are potentially more peripheral to the network [63] unless they are connected to popular others. In semantic networks, total degree centrality may represent the "importance" of a concept or its key concept status. A key concept with high degree centrality is able to activate many other key concepts; thus, it functions as a hot topic's central key concept [54]. Using only the local structure to calculate the degree centrality of a node, this measure alone does not take into consideration the position of the concept within the global structure of the network. Still, the distribution of network statistics, such as node centrality and even node and edge counts, can be an indication of global network properties [61, 64]. In this paper, we employ the weighted version of total degree centrality.

Concept frequency is arguably a more parsimonious metric than popularity (i.e., total degree centrality). However, since we are interested in the semantic structure, focusing on popularity over frequency is appropriate. While frequency alone would reveal some of the key concepts, the semantic linkages (such structures surrounding connective concepts) would remain obscured. Still, a naïve Pearson correlation between the two metrics (weighted degree centrality and frequency) is high (ECB $r \approx 0.77$; Fed $r \approx 0.92$) for both organizations and across all of the periods. A closer inspection reveals significant variance in their relationship and that a log-linear association emerges but only for those concepts having higher than average frequency and degree. In fact, given the heavy positive skewness of weighted degree centrality distributions (i.e., many low values), we inspect a $\log + 1$ transformation of degree centrality which reveals more comparable correlations: ≈ 0.71 and ≈ 0.77 for the ECB and Fed, respectively. The ranked correlations (i.e., Kendall's τ) are even lower (≈ 0.55 and ≈ 0.70, respectively); thus, using concept frequency *in lieu* of weighted degree centrality would reveal only a semi-structural space, having non-trivially different interpretations than the space we employ in this paper.

Betweenness centrality

Betweenness centrality (C_B) is the sum of the proportions of the shortest paths a node lies on for every pair of nodes (out of all shortest paths for each pair). The formulation for unweighted betweenness is:

$$C_\text{B}(i) = \sum_{s \neq i \neq t} \frac{\sigma_{s,t}(i)}{\sigma_{s,t}}$$

where $\sigma_{s,t}$ indicates the count of shortest paths between nodes s and t. For weighted betweenness, the shortest paths are computed using the inverse of the edge weight since heavier edges should warrant greater flow (and hence higher betweenness). We

employ this inversion as most of the edges between concepts are valued (i.e., weighted). More broadly, betweenness centrality represents the frequency with which a particular node is on the geodesic path between any other two nodes in the network [8]. As such, betweenness centrality captures one aspect of a node's position in the graph, thus taking into account the global structure of the network. The effectiveness of this measure is limited by the extent of connectedness in the network. That is, if a network contains many disconnected components, betweenness centrality becomes less of a global measure. However, this is not an issue with our semantic networks because each constitutes a single component. The betweenness centrality of a concept within a semantic network is a direct indicator of its influence [55, 65, 66]. A key concept with high betweenness centrality controls access to other key concepts in the network [10, 62, 67, 68] and thus serves as a gatekeeper between different domains [69]. For semantic networks, it is presumed that a node with high betweenness centrality has a higher likelihood to become activated or activate connections across domains (or topic communities).

Structural roles

By combining popularity and connectivity of concepts in semantic networks, we expect to capture emerging topics within the texts and subtle shifts in formal discourse through the classification of nodes according to their structural roles.

Because the discourses of the Fed and the ECB are highly formal and the resulting networks are complex, looking separately at (1) the top most frequent concepts, (2) the top most central concepts (see example in Table 3), or (3) concepts having the highest betweenness centrality (see example in Table 4) will not be very informative.[3] These top concepts are very similar across the four periods for both organizations and constitute the core issues under discussion.

In order to explore both in-depth and orthogonally informative dimensions of the ECB and the Fed discourses, we characterize these discourses using two distinct measures, building on the manner in which popular and connecting concepts play different roles in the structure and dynamics of semantic networks. Combining the popularity (i.e., total degree centrality) and connectivity (i.e., betweenness centrality) dimensions allows for the identification of four structural roles. This combination positions the concepts within this structural role space.

Table 3 Concepts with the highest total degree centrality in each ECB network

Pre-crisis		Crisis		Post-crisis		Recovery	
Concept	Deg.	Concept	Deg.	Concept	Deg.	Concept	Deg.
ECB	8200	ECB	9800	ECB	11,300	ECB	12,700
European	5600	Market	6200	Financial	6900	Bank	10,400
EU	5200	Central bank	6200	Euro area	6300	European	9400
System	4900	Eurosystem	5900	Market	6200	Financial	8200
Eurosystem	4700	EU	5500	Bank	6000	Include	7500
Euro area	4600	Euro area	5200	Include	5900	Market	7400
Central bank	4500	Operate	4800	System	5800	Monetary	7300
Include	4500	National	4500	EU	5800	EU	7200
Market	4300	Include	4200	Central bank	5300	Central bank	6900
Operate	4000	Increase	4100	Economic	5200	Euro area	6900

Table 4 Concepts with the highest total betweenness centrality in each Fed network

Pre-crisis		Crisis		Post-crisis		Recovery	
Concept	Deg.	Concept	Deg.	Concept	Deg.	Concept	Deg.
Bank	77,942	Fed	250,152	Bank	4,767,507	Bank	106,529
Fed	25,589	Bank	232,961	Fed	2,242,646	Capital	89,368
Rate	24,073	Loan	96,151	Credit	1,273,791	Fed	89,311
Financial	20,868	Credit	73,132	Financial	1,244,069	Financial	59,273
Agency	12,662	Financial	41,794	Loan	675,665	Committee	48,583
President	10,151	Capital	33,451	Agency	552,549	Credit	41,855
Credit	9809	Market	27,397	Institution	452,635	Agency	39,651
Institution	8804	Agency	26,388	Rate	422,863	Loan	32,514
Board	6860	Institution	21,378	Consumer	422,569	Institution	17,825
Loan	6730	Rate	15,927	Mortgage	364,074	Rate	16,707

Correlations among centrality and text-analytic measures

While high correlations among measures can be used to identify obviously important network nodes (e.g., Huang et al. [33]), our structural space approach would be limited if degree and betweenness centralities in semantic networks (including our own) were highly correlated and it would fail to identify subtle roles of network nodes.

Research on correlations between degree and centrality measures in empirical networks and well-known topologies is mixed. Li et al. [70] find moderate Pearson correlations between degree and betweenness centralities (average ≈ 0.6) in real-world networks and significant overlap in top node memberships (for nodes ranked by each degree and betweenness centrality) in Erdős-Rényi (ER) networks [71] and scale-free (SF) networks [61] (>0.95). However, this finding is limited as their simulated networks were of one size and specific densities. Also, Pearson correlations can be misleading as they can be heavily influenced by outliers (e.g., a single node having extremely high betweenness and degree centralities). Lee and Pfeffer [72] show that while ER, Watts-Strogatz (WS) small world networks [73], and SF networks display high average Pearson correlations (for degree and betweenness centralities) across a range of network sizes and densities, these correlations can be poor for networks with low densities or high centralization; these characteristics are found in many empirical networks.

Thus, while correlations between centrality metrics require further investigation, it is fair to argue that highly correlated degree and betweenness centrality metrics would have major implications on their joint interpretation. The semantic networks explored in this particular study exhibit only moderate correlations between degree and betweenness centrality (Kendall's $\tau \leq 0.51$) as shown in column DB of Table 5.[4] Thus, these measures are distinct, yet conceptually related, and can be jointly used to reveal distinct structural roles.

Table 5 Average correlations among text-analytic and semantic network measures

	FT	DF	BF	DT	BT	DB
ECB	0.704 (0.059)	0.553 (0.025)	0.496 (0.016)	0.383 (0.031)	0.391 (0.021)	0.517 (0.045)
Fed	0.867 (0.024)	0.709 (0.023)	0.472 (0.026)	0.679 (0.033)	0.477 (0.024)	0.422 (0.009)

F frequency, *T* tf-idf, *D* degree centrality, *B* betweenness centrality
These correlations are averaged across the four periods. The standard deviations appears just below the averages

Nerghes *et al. Computational Social Networks* (2015) 2:16

Page 10 of 29

One might consider the use of direct text-analytic metrics such as frequency of words (or in our case, frequency of concepts) and the term frequency-inverse document frequency (tf-idf) score.[5] In Table 5, we also report average Kendall's τ-rank correlation coefficients among concept frequency, tf-idf, and the two centrality measures; the correlations are averaged across the four periods for each of the data sets. The correlations between either text-analytic measure and the centrality measures are modest, indicating that the network measures are not substitutable (with text-analytic measures). In fact, the two text-analytic measures exhibit the highest correlations (within each data set) and are more substitutable for one another than for the network measures. So, while a direct analysis through frequency counts or tf-idf values of the documents issued by the ECB and the Fed could potentially reveal noteworthy findings, we argue that semantic network analysis adds a new dimension to this text analysis by revealing deeper insights into discursive structures.

In order to connect the concepts with these structural roles, each concept in the network has been ranked based on its total degree centrality (C_D) and its betweenness centrality (C_B). For these rankings, we first consider the *set* of unique, unordered values X derived from some vector (or bag) of measures X_b. The order set of X is then:

$$X_{\text{ordered}} = \{x_1, \ldots, x_n | x_i \in X; n = |X|; x_1 < \cdots < x_n\}$$

where n is the number of unique measure values. We also define an index set J such that $x_j^{\text{ordered}} | j \in J$ is the jth element of ordered set X_{ordered}. We now have a bijection $X_{\text{ordered}} \to X_{\text{rank}}$. For degree centrality, we replace X_b with the degree centrality measures C_D and obtain X_{ordered} which contains the unique, ordered degree centralities. For each node i, $C_D(i)$, we obtain the degree centrality rank $C_D^{\text{rank}}(i)$:

$$C_D^{\text{rank}}(i) = 100 \cdot \frac{j}{n} | \left(x_j^{\text{ordered}} = C_D(i) \right).$$

The rankings are rescaled (hence normalized) to the [0,100] interval so that we can easily compare rankings across semantic networks. The rankings for betweenness centrality are obtained in a similar fashion (i.e., using C_B for X_b).

In simpler terms, we rank the total degree centrality and betweenness centrality scores for the concepts from each time period network into a normalized range between 0 and 100. Ranking was employed because (a) the networks are of different sizes and densities (hence, we want to be able compare across time periods) and (b) using the raw centrality scores produces less compelling and readable visualizations due to high skewness of the centrality distributions.[6]

We highlight the four structural roles as four quadrants of the four structural space in Fig. 1. The Globally Central (GC) role includes concepts with high degree centrality and high betweenness centrality (high values of $C_D^{\text{rank}} \times C_B^{\text{rank}}$), where C_D^{rank} and C_B^{rank} are the normalized rankings of total degree centrality and betweenness centrality, respectively. A GC concept is a central key concept of a hot topic because not only is it highly connected to other concepts (i.e., popular) but it also serves as a bridge between different parts of the network (i.e., connective).

The Locally Central (LC) role contains concepts with high degree centrality and low betweenness centrality (high values of $C_D^{\text{rank}} \times (100 - C_B^{\text{rank}})$). LC concepts are very popular concepts that do not have a strongly connective role. In other words, an LC concept is the

Fig. 1 The four quadrants of the structural space
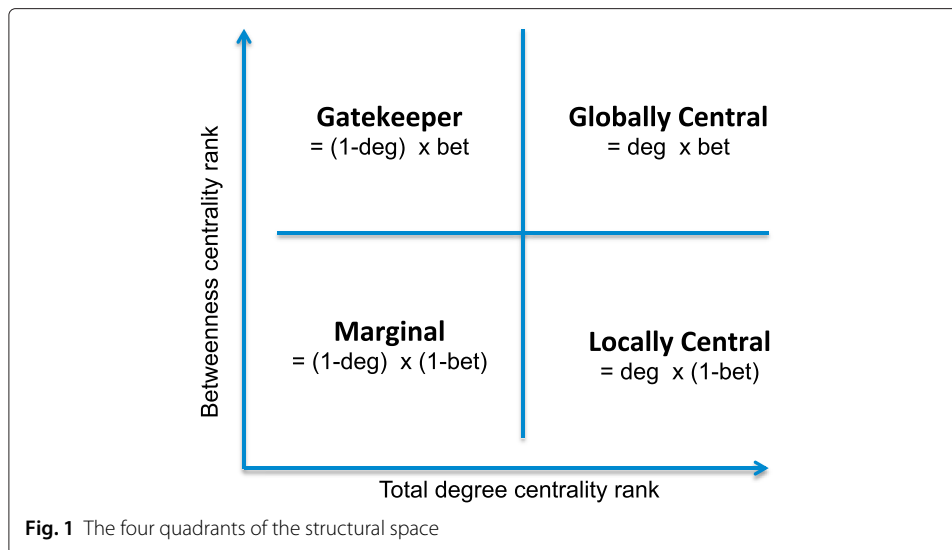
central key concept of a local hot topic because it is highly connected but does not serve as a bridge in the network.

The Gatekeeper (G) role incorporates concepts with low degree centrality and high betweenness centrality (high values of $(100 - C_D^{\mathrm{rank}}) \times C_B^{\mathrm{rank}}$). These types of concepts are highly connective concepts that aren't very popular. A G concept is influential in the network because, although it is not highly connected, it acts as a bridge in the network, linking different themes or topics. Such a concept can mark the emergence of merging themes.

Lastly, the Marginal (M) role includes concepts with low degree centrality and low betweenness centrality (high values of $(100 - C_D^{\mathrm{rank}}) \times (100 - C_B^{\mathrm{rank}})$). M concepts are neither popular nor connective but retain the potential to become emergent concepts and transition into other roles.

Figure 2 illustrates a scatterplot layout example for the four structural roles described above using empirical centrality ranks from one of our subsamples. The darker the red shade of the nodes, the higher embedded these nodes are in the region of the specific structural role.

Alternatively, the structural role scores could have been computed by simply adding the total degree centrality and betweenness centrality score components (i.e., the multiplicands). However, this addition produces inaccurate role mappings. For example, GC concepts become classified also as G due to their high betweenness centrality irrespective of their high total degree centrality. Similarly, M concepts can appear as Gatekeepers due to their extremely low total degree centrality. We find multiplication of the role components to parsimoniously produce distinct role assignments.

After identifying the four roles and the nodes that belong to each role, additional visual dimensions can be added by sizing, shaping, and/or coloring the nodes based on other measures. This approach provides other ways to explore each role individually or the structural space as a whole in terms of distinct or subtle patterns. For example, other network metrics may expose other types of structural roles.

While, in semantic networks, total degree centrality represents the popularity of a concept and betweenness centrality represents the links between two different thematic

Nerghes *et al. Computational Social Networks* (2015) 2:16

Page 12 of 29

Fig. 2 Example of structural roles
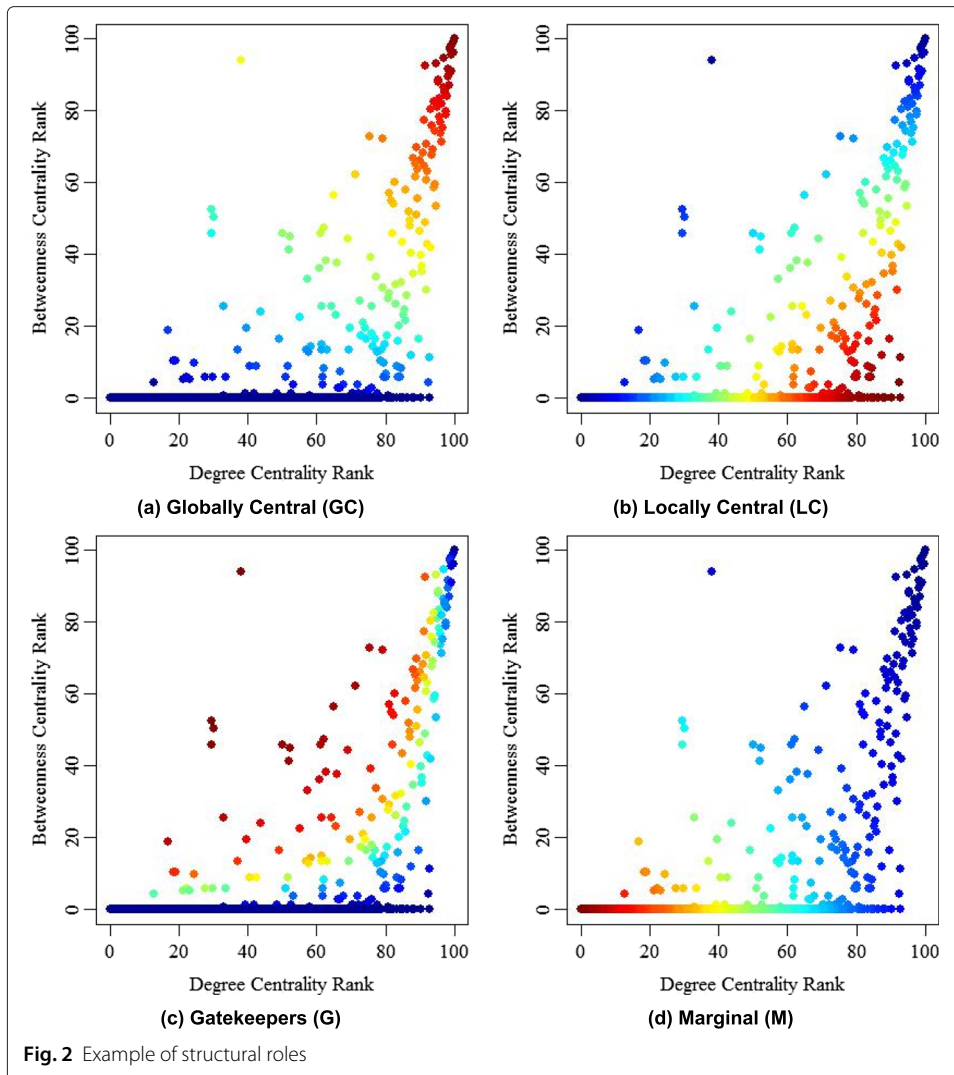
areas, the combination of these two measures has the potential to uncover more subtle structural properties of concepts and thus a set of changes in discourse over time.

Results

Structural roles

For the purposes of this article, we will focus our structural space analysis on a set of selected key concepts within the discourses of the ECB and the Fed and the dynamics of these concepts across four stages of the financial crisis. As presented in Table 6, the first set of concepts have been selected because they reflect the main objectives of the ECB [34] and the Fed [74]. The second set of concepts has been included in the analysis to represent crisis-oriented terminology. While six of the crisis-oriented concepts are similar for both organizations, the other four represent financial instruments employed by the ECB and the Fed during the different stages of the crisis. Below, we describe the acronyms in Table 6:

- *Longer Term Refinancing Operations (LTRO)* provide additional, longer term refinancing to the financial sector;

Nerghes *et al. Computational Social Networks* (2015) 2:16

Page 13 of 29

Table 6 Selected key concepts

ECB	Fed
Main objectives	
Interest rate	Interest rate
Stability	Stable price
Growth	Employment
	Unemployment
Crisis-oriented concepts	
Crisis	Crisis
Debt	Debt
Inflation	Inflation
Lend	Lend
Refinancing	Refinancing
Risk	Risk
LTRO	TALF
MRO	TSLF

- *Main Refinancing Operations (MRO)* serve to drive short-term interest rates, to manage the liquidity situation, and to signal the monetary policy stance in the euro area;
- *Term Asset-Backed Securities Loan Facility (TALF)* was created in 2008 to accommodate the credit needs of consumers and small businesses by facilitating the issuance of asset-backed securities collateralized by loans such as student loans, auto loans, credit card loans, commercial mortgages, and loans guaranteed by the Small Business Administration;
- *Term Securities Lending Facility (TSLF)* is a special lending facility set up by the Federal Reserve in 2008 to loan Treasury securities to primary dealers for 28 days.
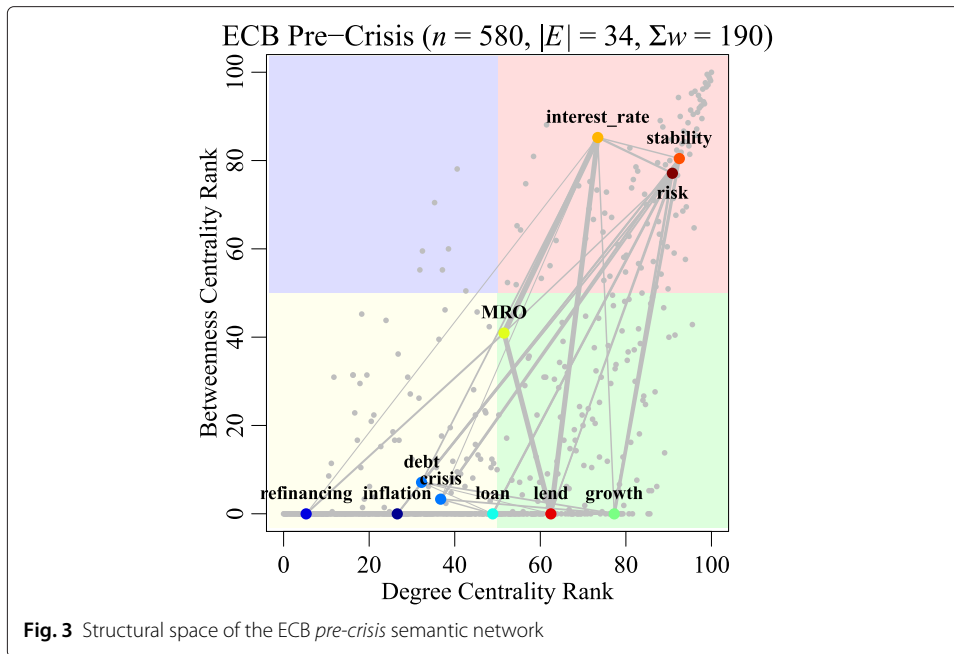
We will begin by reporting our findings for each organization separately by highlighting the observed variations in discourse as it develops across the different phas es of the financial crisis.

First, in Figs. 3, 4, 5, 6, 7, 8, 9, and 10, we plot the structural space positions of all the ECB concepts (with a raw frequency ≥ 10) in the semantic networks of the ECB later followed by those of the Fed. We highlight the selected key concepts by labeling them in the structural space, and we add another visual dimension by coloring the nodes based on their raw frequencies of occurrence in the text data. The color spectrum ranges from dark *blue* (low frequency) to dark *red* (high frequency).
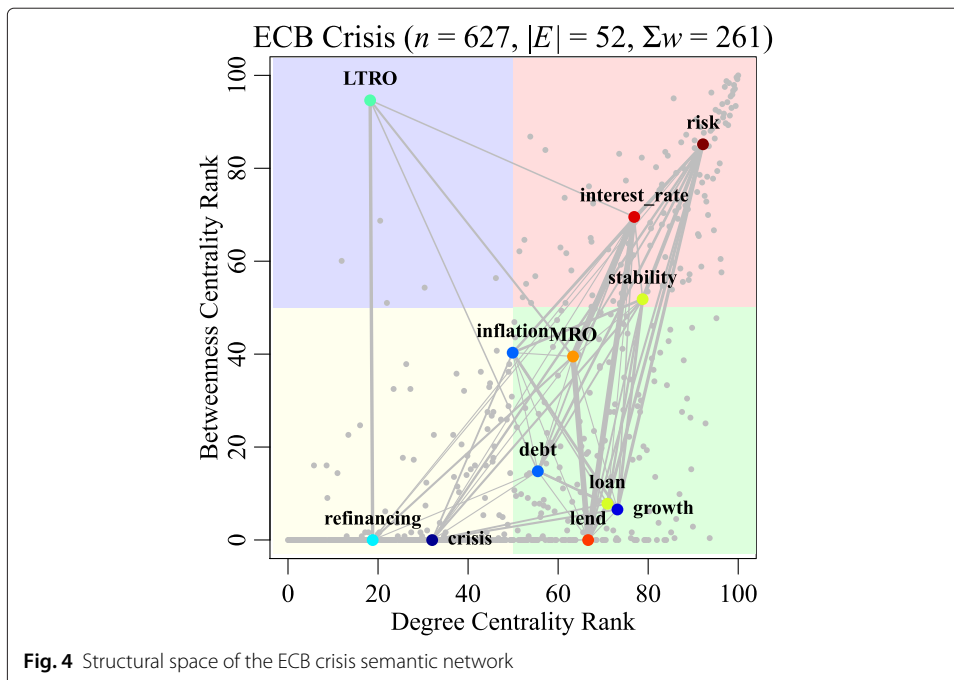
We also display edges among the focal concepts, that is, the subgraph induced by the node set comprising these concepts. The edges are visually weighted and represent the count of co-occurrences (within the one-sentence window) between these focal nodes. We note that these edges do not represent the total activity of the focal concepts but just the activity among themselves (for presentation purposes). Finally, at the top of each graph, we report (1) the number of considered nodes n (i.e., having a raw frequency ≥ 10), (2) the number of distinct edges of the subgraph of key concepts $|E|$, and (3) the sum of the edge weights of that subgraph Σw.
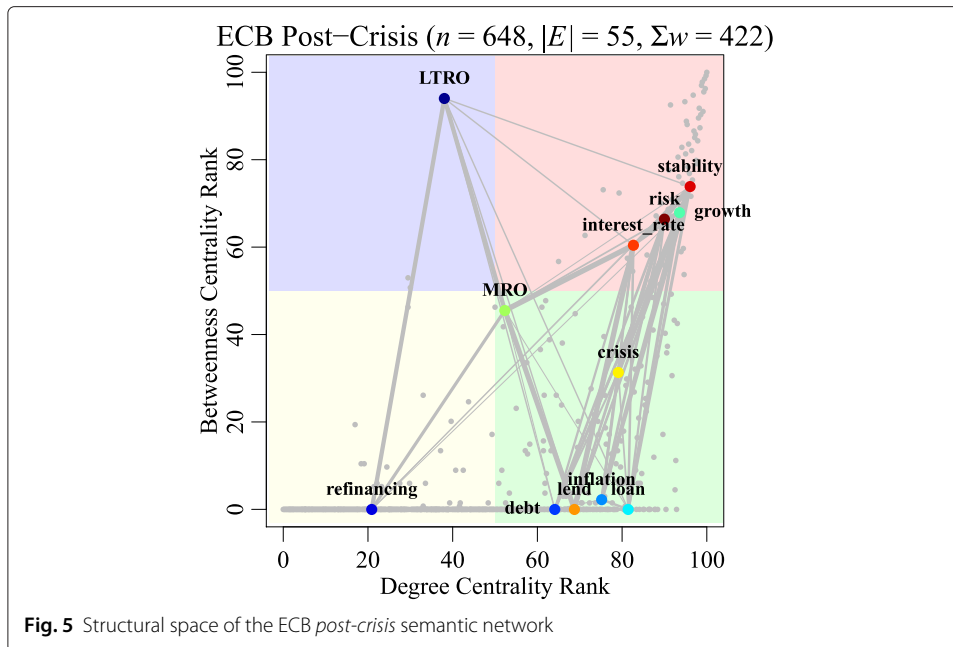
ECB

Figure 3 shows that even in the *pre-crisis* period, before the end of 2007, crisis-oriented key concepts are present in the ECB discourse, some of them having relatively high total
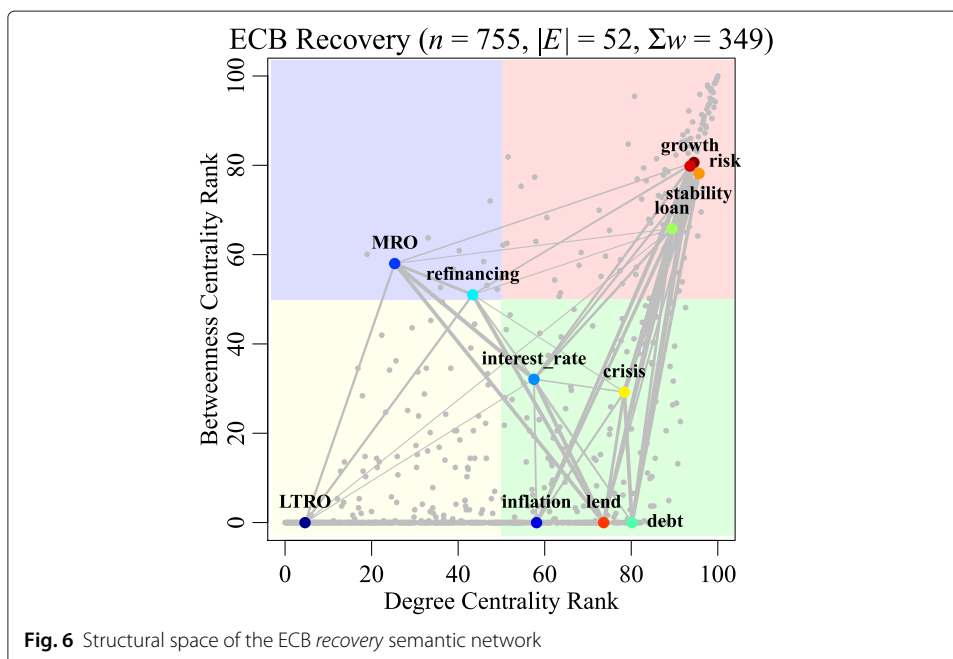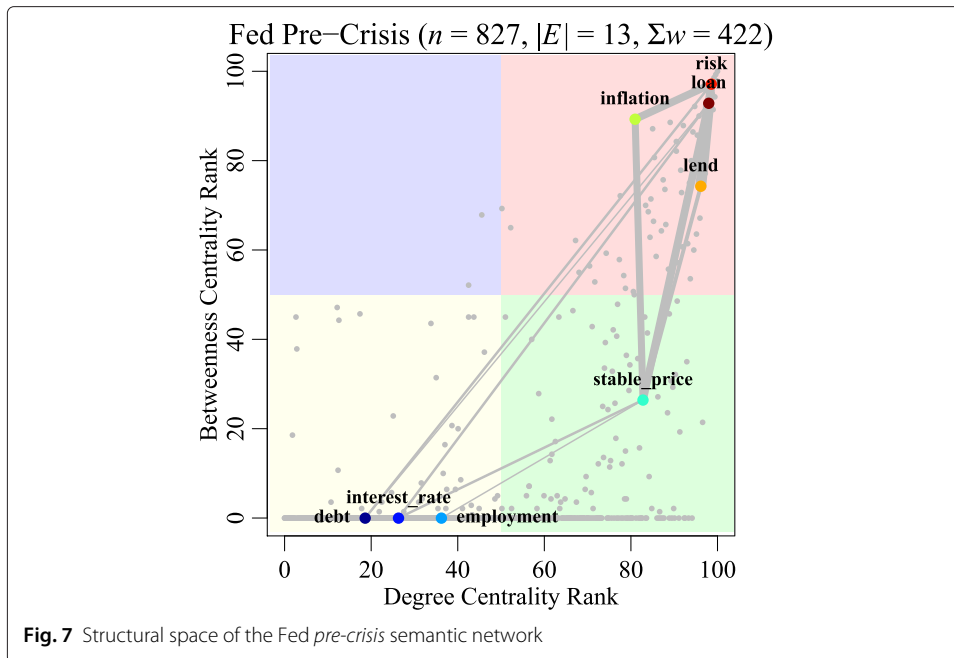
Nerghes *et al. Computational Social Networks* (2015) 2:16

Page 14 of 29

Fig. 3 Structural space of the ECB *pre-crisis* semantic network

degree centrality (i.e., "MRO" and "risk") and being connected to the main objectives of the ECB. The GC position of "risk" as well as the "on-the-fence" position of "MRO" (which borders the LC and the M quadrants) could indicate that some of the ECB's attention was focused on the emerging financial crisis before the end of 2007. We also observe that unlike "interest_rate" and "stability," which are highly ranked GC concepts, "growth" (one of the ECB's main objectives) is a highly ranked LC concept. This indicates that, during the *pre-crisis* period, "growth" was a popular concept but not a very connective one.
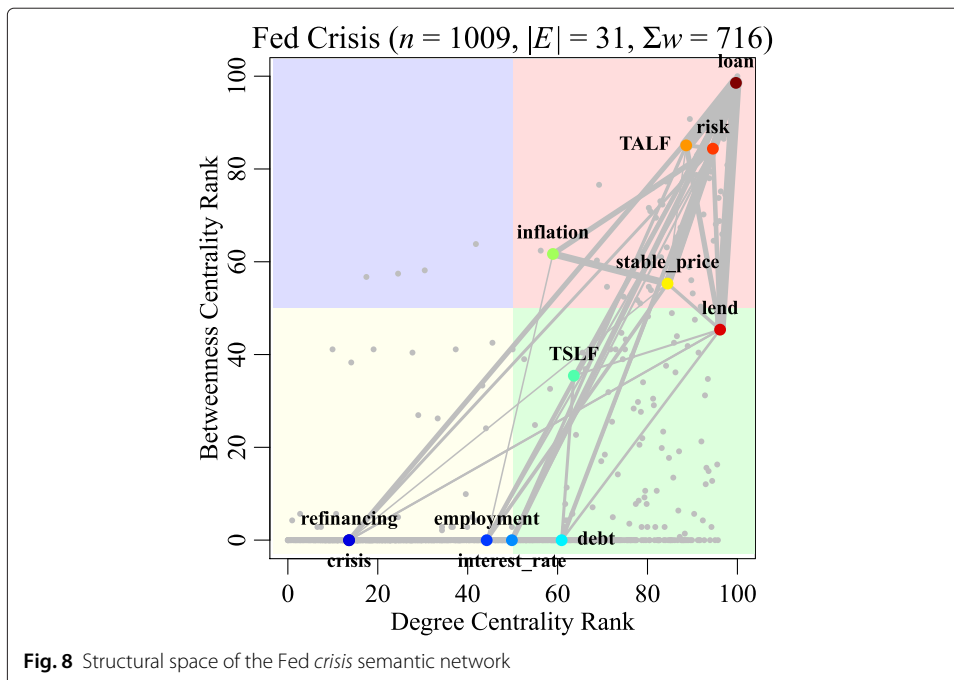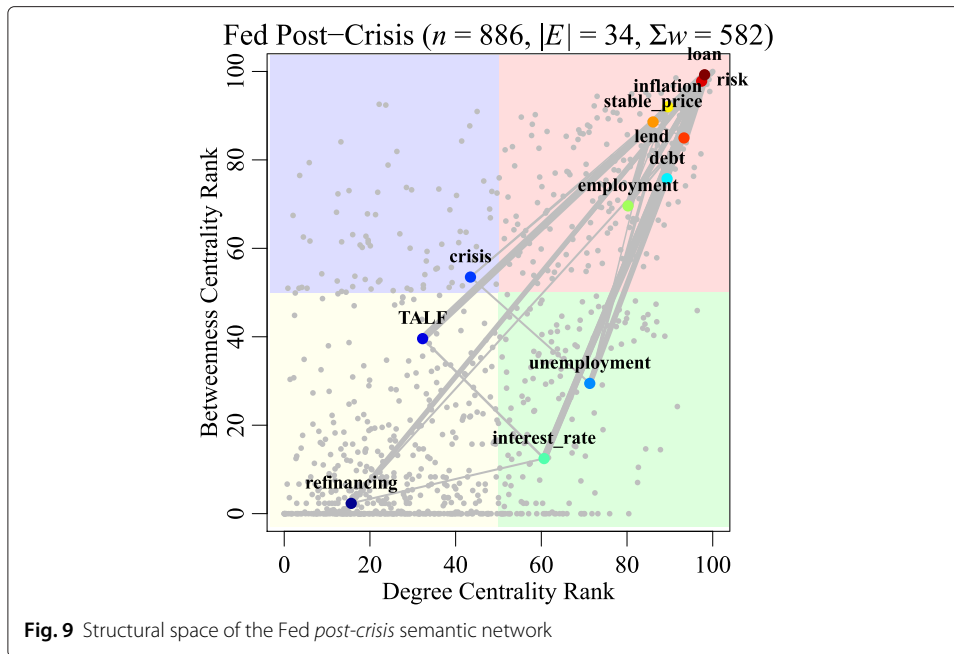

Fig. 4 Structural space of the ECB crisis semantic network

Nerghes *et al. Computational Social Networks* (2015) 2:16

Page 15 of 29

Fig. 5 Structural space of the ECB *post-crisis* semantic network

In the *crisis* period (see Fig. 4), most of the crisis-related key concepts are becoming more prominent. Concepts such as "inflation," "loan," "MRO," and "debt" are ranked higher in total degree centrality and in betweenness centrality than in the previous period, suggesting they became more central and connective of different domains in the discourse of the ECB during the *crisis* period. At the same time, the betweenness and degree centralities of "interest rate" and "stability" noticeably decrease, suggesting once again that the main objectives of the ECB lose rhetorical ground against the full-blown financial crisis. The concept "risk" is even higher ranked in the GC category during the *crisis*
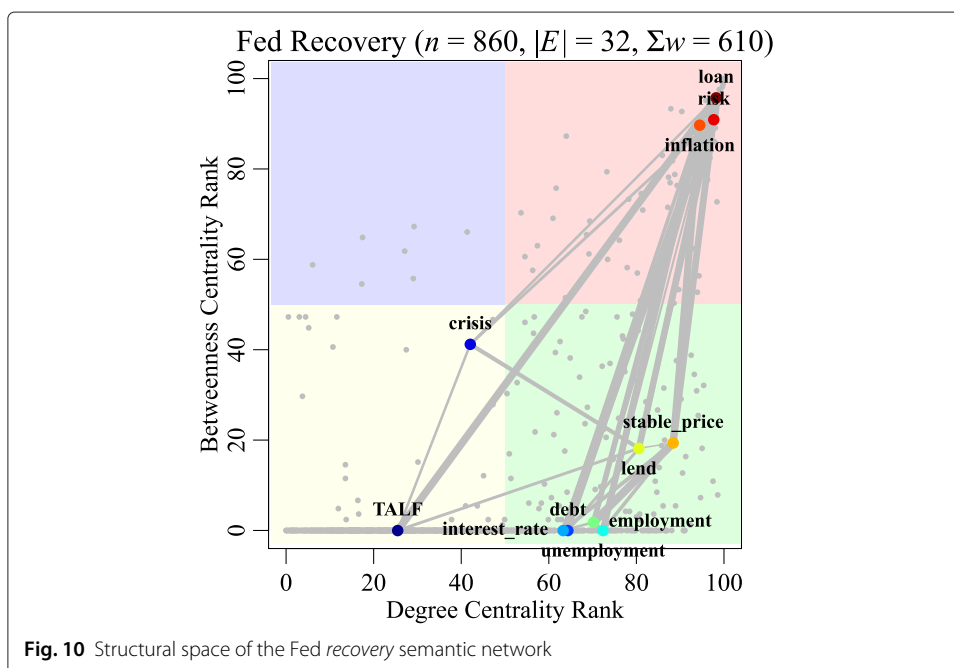

Fig. 6 Structural space of the ECB *recovery* semantic network

Nerghes *et al. Computational Social Networks* (2015) 2:16

Page 16 of 29

Fig. 7 Structural space of the Fed *pre-crisis* semantic network

becoming one of the "hottest" topics of the ECB discourse. We also observe the emergence of "LTRO," a concept that was not present in the *pre-crisis* period. "LTRO" enters the discourse of the ECB as a very highly ranked G concept, indicating that it connected disparate topics during the *crisis*. The link weights also increase in the *crisis* period, highlighting the increased co-activity among these concepts in the ECB press releases during this period. An interesting finding is that "crisis" remains a marginal concept during the *crisis* period. The similar marginal position of "crisis" during the *pre-crisis* as well as the

Fig. 8 Structural space of the Fed *crisis* semantic network

Nerghes *et al. Computational Social Networks* (2015) 2:16

Page 17 of 29

Fig. 9 Structural space of the Fed *post-crisis* semantic network

crisis period raises questions regarding the discursive practices employed by the ECB. Had the ECB avoided the use of such concepts to avoid creating panic among stakeholders? Or had the ECB denied or ignored the existence of the crisis? Or did they perhaps use less value-laden concepts?

Figure 5 plots the structural space of the *post-crisis* semantic network, showing all the main objectives of the ECB in the GC quadrant. While in the *pre-crisis*, only two of the main objectives were in GC positions and in the *crisis* period the centrality of these two concepts decreased, in the *post-crisis* period, all the three main objectives return to being globally central, GC. At the same time, we observe that the betweenness centrality

Fig. 10 Structural space of the Fed *recovery* semantic network

Nerghes *et al. Computational Social Networks* (2015) 2:16

Page 18 of 29

of "crisis" increases, while the betweenness of "debt" and "inflation" decreases significantly. These positional changes suggest that the shift in the ECB discourse could be at least partly explained by its struggle to deal with the aftermath of the crisis, while at the same time refocusing on its core objectives. The link weights show increased activity for "LTRO," "MRO," and "refinancing," lending further evidence to the ECB's resumed focus on the aftermath of the crisis.

Figure 6 plots the structural space of the *recovery* semantic network, revealing significant changes in the discourse of the ECB beyond the crisis. Compared to the previous period, "interest_rate" is now an LC concept. This concept, signifying one of the main objectives of the ECB, maintained a GC position in all the three previous periods analyzed. While "LTRO" suffers a drastic decrease in betweenness centrality (becoming an M concept), "MRO" and "refinancing" become G concepts. We also note the positional change of "loan," moving from the LC quadrant to the GC quadrant. Based on all these structural changes, we argue that the *recovery* period exhibits a clear shift towards a discourse directed at dealing with the aftermath of the crisis. By assessing the width of the links, we see that the positional changes described above are also reflected in the co-occurrence levels. While "MRO" and "refinancing" show increased activity, "LTRO" co-occurs less often with the other key concepts.

As for the graph-level statistics, we observe that the count of nodes (i.e., *n*, the count of non-infrequent concepts) increases almost exactly linearly to the word counts of non-noise words in the collected documents for each period. These word counts are 28,155; 30,991; 33,538; and 42,892 (from *pre-crisis* through *recovery*), and the correlation is 0.997. However, the activity in the focal concept subgraph does not follow suit. Specifically, the edge count increases initially and then stabilizes at ~ 53, and the sum of edge weights peaks at *post-crisis* and then decreases. We surmise that the ECB discourse becomes expansive with the inclusion of additional topics (not identified in this paper). Hence, a naïve analysis using simple, relative frequencies of these key concepts would only diminish their importance. On the other hand, our structural role analysis reveals that some of the concepts associated with the ECB's objectives (here, "stability" and "growth") in fact remain prominent.

Fed

In Fig. 7, we see that, in the *pre-crisis* period for the Fed data, 8 of the 13 selected key concepts are present. Interestingly, two of the Fed's main objectives are marginal concepts, thus having low popularity and low connectivity. The third concept representing the Fed's main objectives, "stable_price," is a popular concept (LC concept) during the *pre-crisis*, but it does not serve as a bridge between other topics discussed. The three highest ranked GC concepts in this period are "inflation," "loan," and "risk," concepts that we categorized as crisis-oriented terminology. With "risk" being the highest ranked concept for this period and with the M and LC positions of the three main objectives of the Fed, we can conclude that the Fed's discourse is not focused on their main objectives but on what appears to be the initial signs of the financial crisis.

In the *crisis* period (see Fig. 8), 12 of the 13 selected key concepts are present. Similar to the previous period analyzed, not all of the main objectives of the Fed are present, and while "employment" maintains its marginal position, "stable_price" moves from the LC quadrant into the GC quadrant, becoming a more popular and a more connective

concept during this period. Unlike the *pre-crisis* period which had "risk" as the top GC concept, in the *crisis* period, the highest ranked GC concept is "loan" (a concept that was already a GC concept in the *pre-crisis* period). We also observe that "Term Asset-Backed Securities Loan Facility" (TALF) and "Term Securities Lending Facility" (TSLF) enter the discourse of the Fed, with "TALF" being a highly popular and connective concept, while "TSLF" is mostly a popular concept. The *crisis* period clearly reveals the discourse of the Fed focusing on the events of the financial crisis, while at the same time being focused on maintaining price stability. Just as in the *pre-crisis* period, "unemployment," the fourth main objective of the Fed, is not present. Based on the node count (*n*) of the *crisis* and the *pre-crisis* periods, we can conclude that the Fed's discourse expands in the *crisis* period. In other words, the Fed has been using more unique concepts than in the previous period. This period also exhibits the most subgraph activity ($\sum w$), indicating the increased co-occurrence of these concepts.

Figure 9 depicts the *post-crisis* period in which 12 of the key concepts are present. During this period, we note that all four of the Fed's main objectives are present, with "employment" shifting from its M position to a GC position. Also, while "stable_price" becomes an even more popular concept, "unemployment" enters the discourse of the Fed directly as a popular concept (LC). It is also worth mentioning the new position of "crisis," an emerging concept in the *crisis* period, which has now become a connective concept (G). The degree centrality and betweenness centrality of "TALF" decrease, placing this concept into the M quadrant.

The *recovery* period presented in Fig. 10 reveals all four main objectives of the Fed as popular but not connective (LC) concepts, while the three prominent GC concepts ("loan," "risk," and "inflation") are all crisis-related terms. Based on these findings, we can argue that while the Fed's focus was shifting towards their main objectives, the aftermath of the crisis was still evident in their discourse.

The graph-level statistics for the Fed also reveal an interesting trajectory. While the *crisis* semantic network has the highest node count (*n*) and the highest subgraph activity ($\sum w$), the *post-crisis* network shows the highest number of links (|*E*|) between the focal concepts. In contrast with our findings for the ECB, where the expansion of their discourse happens gradually over the four periods, the Fed's discourse expands during the *crisis* period but contracts in the following periods. In other words, while the ECB's use of unique concepts grows throughout the four different stages, the Fed's use of unique words expands rapidly between the *pre-crisis* and the *crisis* and decreases thereafter. This finding is also supported by the word counts of the pre-processed, cleaned text documents (without noise words) which are 43,564; 72,919; 53,537; and 50,286 (from *pre-crisis* through *recovery*). The overall Pearson correlation between word counts and concept-node counts here is also high and significant, $r = 0.85$.

MRQAP

As the last part of our analysis, and in light of the findings above, we performed Quadratic Assignment Procedure (QAP) correlations and multiple regressions (MRQAP). These methods compare one or more networks using edges and their weights as data points while controlling for their dependencies such as autocorrelation within the network structure [75]. These methods have been widely used in social network research [76–78] and also applied to research in knowledge and semantic networks [79–81]. This type of

Nerghes *et al. Computational Social Networks* (2015) 2:16

Page 20 of 29

analysis is appropriate for our networks because we are using valued data, and we can characterize each of the four periods as a function of its previous periods for each of the two organizations.

MRQAP is essentially multiple regression predicting the edge weights (including non-edges) of one network from one or more other networks. Typically, the networks (both the dependent and independent ones) are transformed into adjacency matrices, so they contain edge weights as well as zeros for non-edges. These matrices are then elongated into single vectors such that the positions in each vector correspond to the same matrix cell positions. These vectors then serve as the dependent and independent variables in the multiple regression, and an estimated regression coefficient (one for each predictor network) indicates the extent to which an independent network's edge (weights) contributes to the corresponding edge weight in the dependent network.

While the regression coefficients from an MRQAP are identical to those of a least squares regression, their significance scores (i.e., p values) are derived by comparing the estimates against their null distributions obtained from applying the same regression model to a large sample of permutations ($m = 1000$) of the node structure (i.e., node relabelings) thereby controlling for autocorrelation [75]. For example, the inherent popularity of certain concepts is controlled for. The permutations of the predictor matrices represent alternative "worlds" in which the structure remains the same but nodes have been repeatedly re-assigned. The same applies to the computation of a QAP Pearson correlation. The networks are first conformed by node count as the network sizes need to be identical as required by standard regression. For this analysis, the semantic networks of each organization have been conformed through intersection of the node sets, which retains only nodes that are common to all the networks.

ECB

The Pearson correlations reported in Table 7 are moderate to high despite the complexity of the four semantic networks.

Their patterns show what we would nominally expect: proximal time periods bear the most resemblance while those farther apart differ the most. For example, the *pre-crisis* network's correlations diminish with more recent periods. Interestingly, the *post-crisis* and *recovery* periods exhibit more similarity to one another than any other adjacent pairs of periods, suggesting these periods are not as distinct as those other pairs and that *recovery* was likely already underway during the *post-crisis* period. Because the *pre-crisis* and *recovery* networks are the least similar, we can argue that the *recovery* network is a transition phase in the ECB discourse towards a new, possibly hybrid, state and not a resumption of the status quo of the *pre-crisis* period.

We are also interested in the extent to which the first three periods constitute the *recovery* period. Conforming the four periods' networks by the intersection of their common nodes yields (n) 382 nodes per network. In Table 8, we show the results of the MRQAP

Table 7 MRQAP correlations for the ECB semantic networks

	Pre-crisis	Crisis	Post-crisis	Recovery
Pre-crisis	–	0.758*	0.697*	0.593*
Crisis	–	–	0.758*	0.627*
Post-crisis	–	–	–	0.851*

*indicates significance at $p < 0.001$

Table 8 MRQAP coefficients for the ECB semantic networks

Dependent	Intercept	Pre-crisis	Crisis	Post-crisis
Recovery	0.361**	0.037*	−0.071**	0.968**

Adj-R^2 = 0.725
* $p < 0.1$, **$p < 0.001$

regression for predicting the *recovery* network from the earlier periods' networks. As suggested by correlations in Table 7, the *post-crisis* period is the most predictive of *recovery*. Similarly, the *pre-crisis* period adds very little to the *recovery* period; however, the coefficient remains positive indicating a contribution to the similarity. On the other hand, the negative (and significant) coefficient for *crisis'* prediction on *recovery* indicates a slight reversal in the semantic structure from that period. That is, semantic associations of high prominence in *crisis* appear less prominently in *recovery*, controlling for the effects from the other two periods. In other words, the ECB seemed more inclined to focus less on the financial crisis and more on the subsequent recovery, an observation supported by the shifts in the structural roles (Figs. 6, 7, 8, and 9).

In light of the findings of the structural space analysis and the MRQAP, we can conclude that the *recovery* period is a different state in the discourse of the ECB. This new state in their discourse exhibits elements characteristic to the *post-crisis*, moving further away from the *crisis* period yet not reverting to the status quo of the *pre-crisis* period.

Fed

In Table 9, the correlations among the intersection-conformed Fed networks are reported. The intersection node set size here is $n = 356$ nodes.

While the correlations are modestly lower here than for the ECB, their patterns mirror those of the ECB correlations. Specifically, adjacent periods exhibit higher correlations, and those pairs of periods with the greatest temporal distance have the lowest correlations.

We again employ a full MRQAP model predicting the *recovery* period from the prior three periods and report the coefficients in Table 10.

While the pattern of coefficients' magnitudes and valences bears some resemblance to that of the ECB MRQAP model, there are also some striking differences. The coefficient for *pre-crisis* is higher here than it is for the ECB, indicating that the structure of language (and hence the policy focus) in the *recovery* period does not depart as widely from *pre-crisis* in the Fed data as it does for the ECB data. Conversely, the coefficient for *crisis* is negatively higher here than for ECB indicating the language structure in the Fed *recovery* period departs even more from *crisis* period language.

Multidimensional scaling

In Fig. 11, we display multidimensional scalings (MDS) in two dimensions for both the ECB and Fed correlation matrices.[7] The correlations were converted to distances via

Table 9 MRQAP correlations for the Fed semantic networks

	Pre-crisis	Crisis	Post-crisis	Recovery
Pre-crisis	–	0.624*	0.581*	0.499*
Crisis	–	–	0.756*	0.536*
Post-crisis	–	–	–	0.745*

*indicates significance at $p < 0.001$

Table 10 MRQAP coefficients for the Fed semantic networks

| Dependent	Intercept	Pre-crisis	Crisis	Post-crisis
Recovery	0.647*	0.336*	−0.128*	0.817*

Adj-R^2 =0.568
*$p < 0.000$

a $1 - r$ transformation where r is a period pairwise Pearson correlation. In MDS, an eigen-decomposition reduces the dimensionality of a matrix of distances (in our case, the extent of non-correlation) such that a pairwise distance between each pair of data points in the reduced space (in our case, 2-D) is roughly proportional to the original distance between the pair. Thus, distant points (or periods of the ECB and of the FED) in Fig. 11 indicate relatively lower correlation in their semantic networks than proximal period points.

The arrows in the figure indicate the temporal progression from *pre-crisis* all the way through *recovery*. The plotted axes have been aligned across subfigures to allow for comparison. The root mean squared errors between the original correlations (upper triangle) and the Euclidean distances of the MDS coordinates are low (0.006 and 0.008 for the ECB and Fed, respectively), indicating that two dimensions adequately capture the temporal differences of the conformed networks and additional dimensions would not add to the results.

The figures succinctly echo our analysis of the raw correlation matrices: that adjacent periods bear the most similarity (i.e., are closest together in the MDS). Interestingly, the trajectories for both the ECB and Fed are not merely straight lines departing from *pre-crisis*, but instead, there is a curvilinear shape to them, which indicates the language structure partly returning to the status quo. However, we can maintain our earlier assertion that *recovery* is a new state given that the trajectories along one of the dimensional axes (the *x*-axis) proceeds monotonically away from *pre-crisis* (for both the ECB and the Fed discourses). Finally, the higher correlations of the ECB are reflected in the MDS trajectories occupying a smaller area of the space.

Conclusion

The goals of the present article were three fold: (1) we sought to overcome two common challenges in text analysis, namely the size of the text corpora and its formal character, (2)

Fig. 11 MDS of correlations

Nerghes *et al. Computational Social Networks* (2015) 2:16

Page 23 of 29
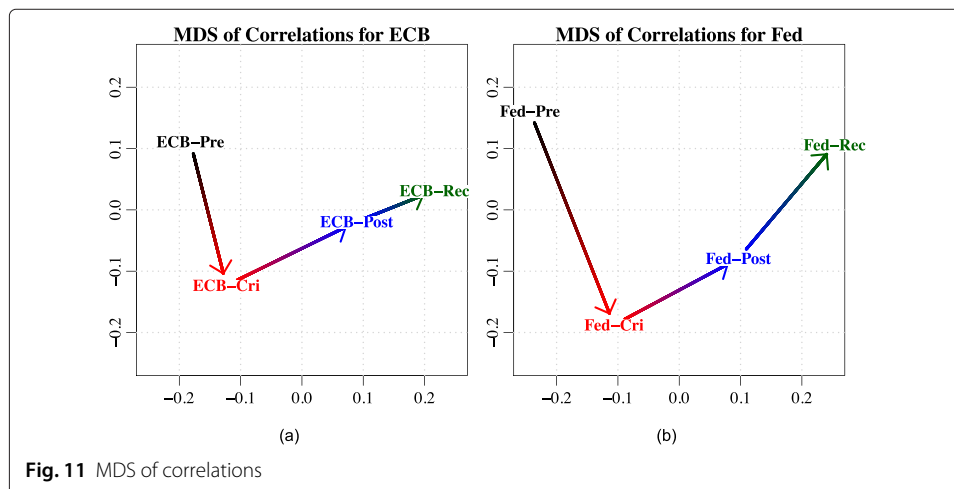
we aimed to explore the benefits of the structural space dimensions, and (3) we wanted to investigate how the discursive practices of the ECB and the Fed have been affected by the recent financial crisis.

The structural space method employed by this study revealed substantial and imperative shifts in the discourse of both the ECB and Fed, demonstrating that it could be a valuable instrument for change detection in formal discourse. As demonstrated, focusing on just the obvious most central concepts in formal discourse does not always reveal the underlying and subtler shifts across the periods investigated. Formal organizational discourse contains repetitive top key concepts, indicative of the obvious and perhaps uninformative central topics of an organization. The structural space analysis proved to be more explanatory regarding the shifts and changes in formal discourse, by combining structural measures and looking beyond the core of the network structure. At the same time, structural roles of key concepts may be good predictors of emerging topics and the dynamics of discursive change.

In recent years, an increasing number of researchers have been focusing on the importance of central banks' communication [38, 42, 82–84]. As key policy instruments and valuable sources of information, the communications of the ECB and the Fed are of critical importance for financial market participants and for society at large. This being said, in times of crisis, the role of these central banks in guiding financial market expectations through communication is particularly important due to higher market uncertainty. While previous research established a link between communications of the ECB and the Fed and their respective impacts on the financial markets [39–41, 44, 45], our focus was directed at uncovering the shifts and adaptations of their discourse in a time of crisis and market volatility.

The structural space dimension of the selected key concepts exposed significant changes in the ECB and the Fed discourses. Below, we summarize our main findings for each organization and each of the examined periods.

ECB pre-crisis: Our analysis showed that crisis-oriented key concepts were already present, suggesting that even before the end of 2007 the ECB's discourse shifted towards crisis terminology, and their focus may have been on the impending financial crisis.

ECB crisis: The crisis-related key concepts detected by our method in the *pre-crisis* period became more prominent in the *crisis* period. We showed that the key concepts associated with the main objectives of the ECB lost ground in front of the effects of the full-blown financial crisis. Also, we noted the emergence of "longer term refinancing operations" (LTRO) as a highly ranked connective (G) concept and similarly the increase in betweenness for "main refinancing operations" (MRO). These changes denote the focus of the ECB on refinancing operations during the *crisis* period. Interestingly, the "crisis" concept was not a highly connective nor a popular concept during this period. The marginal position of this concept in both the *pre-crisis* period and the *crisis* period could denote an intentional attempt of the ECB to minimize panic reactions among the stakeholders, or it could be explained by a narrow focus of the ECB's discourse towards the overwhelming market defaults and not towards the crisis as a whole.

ECB post-crisis: This period revealed the ECB's discourse in a distinct state, where all the main objectives of the ECB are in a prominent position, while "crisis" itself became a more connective concept. The changes observed in this period point towards a focus of the ECB's discourse towards dealing with the aftermath of the crisis.

Nerghes *et al. Computational Social Networks* (2015) 2:16

Page 24 of 29

ECB recovery: While in the *crisis* and the *post-crisis* periods "LTRO" is the highest ranked G concept, during the *recovery* period, it suffers a drastic decrease in betweenness centrality ranking, appearing now as a marginal concept (M). At the same time, we show "loan" becoming more popular, and "MRO" and "refinancing" becoming more connective. These structural changes in the semantic structure show the shift towards dealing with the aftermath of the crisis more clearly than in the previous period.

Finally, our findings revealed that by the end of 2013, the discourse of the ECB had in no way returned to the *pre-crisis* levels but perhaps advanced to a new state altogether. This "new state" could be explained by the role of the ECB in dealing with the aftermath of the financial crisis. Also, during the *recovery* period, the ECB seemed to focus less on the financial crisis and more on the subsequent recovery process. This particular finding was supported by the structural space analysis as well as the MRQAP coefficients and multidimensional scaling.

Fed pre-crisis: The *pre-crisis* period exposes the Fed discourse as having a focus on such terms as "risk," "loan," and "inflation." At the same time, we observe that the main objectives of the Fed are not central to its discourse in this period. With "employment" and "interest_rate" as marginal and emerging concepts and "stable_price" as an LC (popular) concept, we can argue that the focus of the Fed seems to already be aimed at the impending financial crisis and not towards their main objectives.

Fed crisis: During the *crisis* period, the Fed discourse expands considerably, and we also note the increased activity in the induced subgraph. The focus on crisis-oriented terminology becomes even clearer during this period, with the same main objectives present. The prominent positions held by even more crisis-oriented concepts lead us to conclude that the Fed discourse during the *crisis* period has been mostly focused on dealing with the events of the financial markets, while at the same time being focused on maintaining price stability.

Fed post-crisis: At this stage, the Fed's discourse seems to regain a focus oriented towards their main objectives, but the crisis terminology remains prominent at the same time. Interestingly, "crisis," a concept that has been only marginal in the previous period, now becomes a connective concept in their discourse. These findings suggest that while the focus of the Fed was still very much aimed at the effects of the crisis, they also returned to a state in which their main objectives became more popular and connective in their discourse.

Fed recovery: During the final period analyzed, we find the Fed's discourse in an unexpected state, in which the main objective concepts are no longer in the GC quadrant, but instead, they are highly ranked LC concepts. Here, these concepts are popular in the Fed's discourse, but they do not serve a connective role between topics being discussed. In light of these findings, and based on the high GC ranks of "loan," "risk," and "inflation," we can postulate that the Fed is acknowledging the (still) precarious state of the financial system, while at the same time dealing with the repercussions of the crisis.

The Fed data MRQAP analysis results are relatively similar to the ones of the ECB, but they also reveal some noteworthy differences. While the general trend is comparable to the one of the ECB data, the *recovery* period discourse of the Fed moves even further away from the *crisis* period but appears more similar to the *pre-crisis* period discourse. Thus, while the ECB's *recovery* period discourse enters a new state which seems to retain

Nerghes *et al. Computational Social Networks* (2015) 2:16

Page 25 of 29

aspects of the *crisis* period but also returning slightly to the *pre-crisis* status quo, these changes in discourse are even more pronounced in the Fed data set.

In sum, the structural space approach has exposed key findings regarding the subtle shifts in the discursive practices of the ECB and the Fed throughout the different phases of the financial crisis. One of the most striking findings is that unlike the ECB, the Fed has not been entirely focused on all of its main objectives with the exception of the *post-crisis* period. While the three main objectives of the ECB are present in all the four periods, all of the Fed's main objectives are only present in the *post-crisis* and *recovery* periods. During the *recovery* period, all of the Fed's main objectives are present, but they are all positioned in the LC quadrant suggesting that these concepts are only locally central and thus not connective between various topics present in the Fed's discourse.

Furthermore, there is greater presence and prominence in positions of certain crisis concepts in the Fed than in the ECB. That is, more Fed crisis concepts emerged to achieve GC positions in *crisis* and *post-crisis* than the crisis concepts of the ECB. In particular, "TALF" emerged as a GC concept (*crisis*) as did "lend" and "debt" (*post-crisis*), suggesting a change in the focus of the Fed between *crisis* and *post-crisis*. Even in *pre-crisis* and *recovery*, we observe the presence of most of the Fed's crisis concepts as well as GC prominence of several of them. The prominence of crisis terms in these periods, in addition to others, points to the Fed's inclination to alert consumers not only about the actual but also a potential crisis. All these findings indicate the Fed's concerns (more so than the ECB's) being oriented towards elements of the financial crisis with either expectation or caution (in the case of *pre-crisis*) or as a reaction to the present crisis (*crisis*) or in a reparative capacity throughout the aftermath (*post-crisis* and *recovery*). This observation coincides with our earlier claim that the Fed exhibits a greater attentiveness to the crisis based on the summary network statistics alone (see Table 2).

Even if the focus on their main objectives differs between the ECB and the Fed, one similarity worth noting is that the "crisis" concept does not occupy a prominent position in either of their discourses until the *post-crisis* period. Whether this finding is influenced by an organizational strategy meant to prevent further panics in the financial markets or simply a delayed acknowledgment of the crisis from the two organizations, this finding warrants further investigation into the possibility that other less value-laden concepts have been used by the ECB and the Fed to describe the events that were affecting the financial system.

Although the method we have employed in this study revealed important findings, one of its primary limitations stems from the general limitations of semantic network analysis. The process of transforming textual data into networks of concepts (or words) implies a series of coding choices which can greatly impact the results of the analysis. That is, the techniques used when preprocessing the raw text (e.g., removing noise words and removing numbers), the identification of nodes to be included in the network, and/or the parameters used for the creation of links (e.g., window size and/or stop unit) can strongly impact the structure of the resulting network. As such, these coding choices should be closely aligned to the objectives of the researcher and should be chosen with care.

While our method of classifying nodes into one of the four structural roles was used to highlight only a handful of key concepts, the classification may easily be broadened to identify lists of top concepts (e.g., top ten) within each of the roles. This enumeration of the roles offers a more complete depiction of the roles and their evolution.

Nerghes *et al. Computational Social Networks* (2015) 2:16

Page 26 of 29

Our naïve treatment of weighted degree centrality, while typical in network research, raises some concerns. Specifically, weights and the number of distinct ties ought to be considered separately as the same total degree centrality score of a node can arise from starkly different egocentric structures. While the exploration of this issue is beyond the scope of this paper, we hope (and expect) that future research will improve the use of weighted degree centrality in semantic and social network analysis.

As for the complex structure of our semantic networks, some diagnostic tests reveal that they are only mildly small-world and not at all scale-free, contrary to the findings of other work. Still, further investigation (outside the scope of this paper) would be required to determine if these inconsistencies are due to the types of organizational semantic networks inferred in our study or the exact nature of semantic network extraction or simply that semantic networks can vary widely in their topologies.

As for metric comparisons with other research, our within-network correlations for our two centrality measures were modest, echoing some other findings, e.g., [72, 85], and low enough to permit the identification of outliers in structural space, particularly in the gatekeeper (G) and locally central (LC) roles. Also, the differences in the correlations between the two data sets may relate to structural differences between the discourses of the ECB and Fed and warrant further investigation.

Furthermore, the direct text-analytical metrics cannot be used in lieu of the network measures due to lack of high correlation. In fact, semantic network analysis in combination with the structural space approach outperforms more direct text-analytical approaches (i.e., frequencies or tf-idf). For example, the relevant G role of "LTRO" remains obscured under direct text analysis which shows high percentiles/rankings for tf-idf (0.98) and medium-high frequency (0.73) but omits its low popularity/degree (0.15) and high connectiveness/betweenness (0.98).[8] Thus, a text-analytic approach would fail to uncover the highly connective role of "LTRO" and other concepts which bridge prominent topics and themes of the ECB discourse.

Our use of centrality ranks as opposed to actual centrality scores warrants additional, future inquiry. We suspect that in order to compare them more precisely across networks of varying sizes and densities, tighter controls must be exerted. We envision the development of highly robust comparative indices that account for both the relative or ranked centrality score as well as the absolute score.

To conclude, we can argue that our approach proved beneficial for the analysis of large corpora of formal organizational discourse. We anticipate our noteworthy results to open new avenues for semantic network research dealing with formal discourses beyond the context of the financial crisis.

Endnotes

[1] For example, do more frequent yet briefer press releases (such as the Fed's) affect public and organizational perception differently?

[2] Most empirical social network data exhibit much lower densities.

[3] For conciseness, the said tables display only one measure per organizational data source; however, the problem of repetitive highly ranked concepts persists for betweenness centralities of the ECB networks and degree centralities of the Fed networks.

[4] Kendall's τ-ranked correlation is preferred here in order to mitigate the impact of extreme points (outliers).

Nerghes *et al. Computational Social Networks* (2015) 2:16

Page 27 of 29

[5]Tf-idf reflects the importance of a word in a collection or corpus, and it is often used as a weighting factor in information retrieval and text mining. The tf-idf value increases proportionally to the number of times a word appears in the document, but is offset by the frequency of the word in the corpus.

[6]Alternatively, we could have employed normalized centrality scores. However, these exhibit the same skewness and still require transformation. Our approach is mathematically similar to using ranks of normalized scores.

[7]MDS is also known as Principal Coordinate Analysis.

[8]While the tf-idf here coincides with betweenness, the mild correlation between the two measures in Table 5 reveals this is not always the case.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

The study has been designed by AN. Furthermore, AN and J-SL collected and analyzed the data, and AN, J-SL, PG, and IH have contributed to the writing of this paper. AN, J-SL, PG, and IH also contributed to the conceptual development of the methods employed in the paper. AN, J-SL, PG, and IH read and approved the final manuscript.

Author details

[1]Department of Organization Sciences, VU University, Amsterdam, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands. [2]CSTM, University of Twente, Drienerlolaan 5, 7500AE Enschede, The Netherlands.

References

1. Loewenstein, J, Ocasio, W, Jones, C: Vocabularies and vocabulary structure: a new approach linking categories, practices, and institutions. Acad. Manage. Ann. **6**(1), 41–86 (2012)
2. Irvine, JT: Formality and informality in communicative events. Am. Anthropologist. **81**(4), 773–790 (1979)
3. Alvesson, M, Karreman, D: Varieties of discourse: on the study of organizations through discourse analysis. Hum. Relat. **53**(9), 1125–1149 (2000)
4. Grant, D, Hardy, C: Introduction: struggles with organizational discourse. Organ. Stud. **25**(1), 5–13 (2004). doi:10.1177/0170840604038173
5. Leydesdorff, L, Welbers, K: The semantic mapping of words and co-words in contexts. J. Informetrics. **5**(3), 469–475 (2011). doi:10.1016/j.joi.2011.01.008
6. Doerfel, ML, Barnett, GA: A semantic network analysis of the International Communication Association. Hum. Commun. Res. **25**(4), 589–603 (1999). doi:10.1111/j.1468-2958.1999.tb00463.x
7. Shim, J, Park, C, Wilding, M: Identifying policy frames through semantic network analysis: an examination of nuclear energy policy across six countries. Policy Sci. **48**(1), 1–33 (2015)
8. Danowski, JA: Analyzing change over time in organizations' publics with a semantic network include list: an illustration with Facebook. In: ASONAM, IEEE, pp. 954–959, Washington DC, USA, (2012)
9. Carley, KM, Kaufer, DS: Semantic connectivity: an approach for analyzing symbols in semantic networks. Commun. Theory. **3**(3), 183–213 (1993)
10. Grebitus, C, Bruhn, M: Analyzing semantic networks of pork quality by means of concept mapping. Food Quality and Preference. **19**(1), 86–96 (2008). doi:10.1016/j.foodqual.2007.07.007
11. Kleinnijenhuis, J, van Hoof, AMJ, Oegema, D, de Ridder, JA: A test of rivaling approaches to explain news effects: news on issue positions of parties, real-world developments, support and criticism, and success and failure. J. Commun. **57**(2), 366–384 (2007). doi:10.1111/j.1460-2466.2007.00347.x
12. Kim, L: Denotation and connotation in public representation: semantic network analysis of Hwang supporters' internet dialogues. Public Underst. Sci. **22**(3), 335–350 (2011). doi:10.1177/0963662511401784
13. Leydesdorff, L, Hellsten, I: Metaphors and diaphors in science communication: mapping the case of stem cell research. Sci. Commun. **27**(1), 64–99 (2005). doi:10.1177/1075547005278346
14. Leydesdorff, L, Hellsten, I: Measuring the meaning of words in contexts: an automated analysis of controversies about 'Monarch butterflies,' 'Frankenfoods,' and 'stem cells'. Scientometrics. **67**(2), 231–258 (2006). doi:10.1007/s11192-006-0096-y
15. Nerghes, A, Hellsten, I, Groenewegen, P: A toxic crisis: metaphorizing the financial Crisis. Int. J. Commun. **9**, 106–132 (2015)
16. Borge-Holthoefer, J, Arenas, A: Semantic networks: structure and dynamics. Entropy. **12**(5), 1264–1302 (2010). doi:10.3390/e12051264
17. Bales, ME, Johnson, SB: Graph theoretic modeling of large-scale semantic networks. J. Biomed. Inform. **39**(4), 451–64 (2006). doi:10.1016/j.jbi.2005.10.007
18. Steyvers, M, Tenenbaum, JB: The large-scale structure of semantic networks: statistical analyses and a model of semantic growth. Cogn. Sci. **29**(1), 41–78 (2005). doi:10.1207/s15516709cog2901_3

Nerghes *et al. Computational Social Networks* (2015) 2:16

Page 28 of 29

19. Postma, E, Wiesman, F, van den Herik, H: Small-world semantic networks. In: Proceedings of the Twelfth Belgium-Netherlands Artificial Intelligence Conference, pp. 191–198, Tilburg University, Tilburg, The Netherlands, (2000)

20. Ferrer I Cancho, R, Solé, RV: The small world of human language. Proc. Biol. Sci. **268**(1482), 2261–5 (2001). doi:10.1098/rspb.2001.1800

21. Solé, RV, Corominas-Murtra, B, Valverde, S, Steels, L: Language networks: their structure, function, and evolution. Complexity. **15**(6), 20–26 (2010)

22. Nerghes, A, Lee, J-S, Groenewegen, P: EU regulators: a structure and content analysis. XXXIII Sunbelt Social Networks Conference of INSNA, Hamburg, Germany (2013)

23. Nerghes, A, Lee, J-S, Groenewegen, P: Market player or regulator? A semantic network analysis of the shifting roles of the ECB during the financial crisis. XXXIV Sunbelt Social Networks Conference of INSNA, St. Pete Beach, FL, USA (2014)

24. Nerghes, A, Lee, J-S, Groenewegen, P, Hellsten, I: The shifting discourse of the European Central Bank: exploring structural space in semantic networks. In: Yetongnon, K, Dipanda, A (eds.) Proceedings of SITIS, pp. 447–455. IEEE Computer Society, Marrakesh, Morocco, (2014). doi:10.1109/SITIS.2014.13

25. Burt, R: Structural holes: the social structure of competition. Harvard University Press, Cambridge, MA (2009)

26. Burt, RS: Cohesion versus structural equivalence as a basis for network subgroups. Sociol. Methods & Res. **7**(2), 189–212 (1978)

27. Burt, R: Detecting role equivalence. Soc. Netw. **12**(1), 83–9 (1990)

28. Boyd, JP, Everett, MG: Relations, residuals, regular interiors, and relative regular equivalence. Soc. Netw. **21**(2), 147–165 (1999)

29. Borgatti, SP, Everett, MG: Notions of position in social network analysis. Sociol. Methodol. **22**(1), 1–35 (1992)

30. White, HC, Boorman, SA, Breiger, RL: Social structure from multiple networks. I. Blockmodels of roles and positions. Am. J. Sociol. **81**(4), 730–780 (1976)

31. Anderson, CJ, Wasserman, S, Faust, K: Building stochastic blockmodels. Soc. Netw. **14**(1), 137–161 (1992)

32. Boorman, SA, White, HC: Social structure from multiple networks. II. Role structures. Am. J. Sociol. **81**(6), 1384–1446 (1976)

33. Huang, S, Lv, T, Zhang, X, Yang, Y, Zheng, W, Wen, C: Identifying node role in social network based on multiple indicators. PloS one. **9**(8), 103733 (2014). doi:10.1371/journal.pone.0103733

34. Treaty, EU Amsterdam: Treaty of Amsterdam amending the treaty on European Union, the treaties establishing the European Communities and certain related acts. European Communities, Office for Official Publications, Luxembourg (1997)

35. Bernanke, B, Reinhart, V, Sack, B: Monetary policy alternatives at the zero bound: an empirical assessment. Brookings Papers Econ. Act. **2**, 1–100 (2004)

36. Gürkaynak, RS, Sack, B, Swanson, E: Do actions speak louder than words? The response of asset prices to monetary policy actions and statements. Int. J. Cent. Bank. **1**(1), 55–93 (2005)

37. Gurkaynak, RS, Sack, B, Swanson, E: The sensitivity of long-term interest rates to economic news: evidence and implications for macroeconomic models. Am. Econ. Rev. **95**(1), 425–436 (2005)

38. Kohn, DL, Sack, BP: Central bank talk: does it matter and why? In: Macroeconomics, Monetary Policy, and Financial Stability, pp. 175–206. Bank of Canada, Ottawa, Ontario, (2003)

39. Jansen, D, Haan, J: Is a word to the wise indeed enough? ECB statements and the predictability of interest rate decisions. Technical Report 75, De Nederlandsche Bank, Amsterdam (2005)

40. Jansen, DJ, Haan, J: Were verbal efforts to support the euro effective? A high-frequency analysis of ECB statements. Eur. J. Pol. Econ. **23**(1), 245–259 (2005)

41. Haan, J: The effect of ECB communication on interest rates: an assessment. Rev. Int. Org. **3**(4), 375–398 (2008). doi:10.1007/s11558-008-9048-z

42. Rosa, C, Verga, G: Is ECB communication effective? Centre for Economic Performance Discussion Paper No. 0682 (2005). doi:10.2139/ssrn.621641

43. The Federal Reserve Board: Mission (2014). http://www.federalreserve.gov/aboutthefed/mission.htm. Accessed 16 April 2014

44. Hayo, B, Neuenkirch, M: Do Federal Reserve communications help predict federal funds target rate decisions? Journal of Macroeconomics. **32**(4), 1014–1024 (2010)

45. Hayo, B, Kutan, AM, Neuenkirch, M: Financial market reaction to Federal Reserve communications: does the crisis make a difference? **8**, 1–18 (2012). MAGKS Joint Discussion Paper Series in Economics, 08-2008, Marburg. doi:10.2139/ssrn.1155481

46. Sahuc, J, Smets, F: Differences in interest rate policy at the ECB and the Fed: an investigation with a medium-scale DSGE model? Journal of Money, Credit and Banking. **40**(2-3), 505–521 (2008). doi:10.1111/j.1538-4616.2008.00124.x

47. Connolly, E, Kohler, M: News and interest rate expectations: a study of six central banks. In: The Future of Inflation Targeting, ed. Christopher Kent and Simon Guttman, pp. 108–34. Reserve Bank of Australia, Sydney, (2004)

48. Chirinko, RS, Curran, C: Greenspan Shrugs: Formal pronouncements , bond market volatility, and central bank communication. In: Proceedings of the American Economic Association Meetings on January, (2006)

49. Diesner, J: From texts to networks: detecting and managing the impact of methodological choices for extracting network data from text data. Künstliche Intelligenz. **27**(1), 75–78 (2013). doi:10.1007/s13218-012-0225-0

50. Carley, KM, Columbus, D, Landwehr, P: Automap user's guide 2013. Technical report, CMU-ISR-CASOS, Pittsburgh, PA (2013)

51. Carley, KM, Palmquist, M: Extracting, representing, and analyzing mental models*. Soc. Forces. **70**(3), 601–636 (1992)

52. Danowski, J: Network analysis of message content. In: Richards, WD, Barnett, GA (eds.) Progress in Communication Science, pp. 197–222. Ablex, Norwood, NJ, (1993). Chap. XII

53. Popping, R: Knowledge graphs and network text analysis. Soc. Sci. Inf. **42**(1), 91–106 (2003). doi:10.1177/0539018403042001798

Nerghes *et al. Computational Social Networks* (2015) 2:16

Page 29 of 29

54. Diesner, J: Uncovering and managing the impact of methodological choices for the computational construction of socio-technical networks from texts. PhD thesis, Carnegie Mellon University (2012)
55. Wasserman, S, Faust, K: Social network analysis: methods and applications, Vol. 24. Cambridge University Press, Cambridge (1994)
56. Freeman, L: Centrality in social networks conceptual clarification. Soc. Netw. **1**(3), 215–239 (1979)
57. Bonacich, P: Power and centrality: a family of measures. Am. J. Sociol. **92**(5), 1170–1182 (1987)
58. Freeman, L: A set of measures of centrality based on betweenness. Sociometry. **40**(1), 35–41 (1977)
59. Wei, T-H: The algebraic foundations of ranking theory (1952)
60. Berge, C: Théorie des Graphes et Ses Applications. Dunod, Paris, France (1958)
61. Barabási, A-L, Albert, R: Emergence of scaling in random networks. Science. **286**(5439), 509–512 (1999)
62. Henderson, G, Iacobucci, D, Calder, B: Brand diagnostics: mapping branding effects using consumer associative networks. Eur. J. Oper. Res. **111**, 306–327 (1998)
63. Iacobucci, D, Henderson, G, Marcati, A, Chang, J: Network analyses of brand switching behavior. Int. J. Res. Mark. **13**, 415–429 (1996)
64. Lee, J-S, Pfeffer, J: Robustness of network centrality metrics in the context of digital communication data. In: Proceedings of HICSS, pp. 1798–1807. IEEE, Kauai, HI, (2015). doi:10.1109/HICSS.2015.217
65. Hill, V, Carley, KM: An approach to identifying consensus in a subfield: the case of organizational culture. Poetics. **27**(1), 1–30 (1999)
66. Hooper, CJ, Marie, N, Kalampokis, E: Dissecting the butterfly: representation of disciplines publishing at the Web Science Conference series. In: WebSci, pp. 197–200. ACM, New York, NY, (2012). doi:10.1145/2380718.2380737
67. Brandes, U, Corman, SR: Visual unrolling of network evolution and the analysis of dynamic discourse†. Inf. Vis. **2**(1), 40–50 (2003). doi:10.1057/palgrave.ivs.9500037
68. Hulst, RC: Introduction to social network analysis (SNA) as an investigative tool. Trends Org. Crime. **12**(2), 101–121 (2008). doi:10.1007/s12117-008-9057-6
69. Gloor, PA, Krauss, J: Web science 2.0: identifying trends through semantic social network analysis. In: Proceedings of CSE, pp. 215–222. IEEE, Vancouver, BC, (2009). doi:10.1109/CSE.2009.186
70. Li, C, Li, Q, Van Mieghem, P, Stanley, HE, Wang, H: Correlation between centrality metrics and their application to the opinion model. Eur. Phys. J. B. **88**(65), 1–13 (2015). doi:10.1140/epjb/e2015-50671-y
71. Erdős, P, Rényi, A: On random graphs. Publ. Math. (Debrecen). **6**, 290–297 (1959)
72. Lee, J-S, Pfeffer, J: Estimating centrality statistics for complete and sampled networks: some approaches and complications. In: Proceedings of HICSS, pp. 1686–1695. IEEE, Kauai, HI, (2015). doi:10.1109/HICSS.2015.203
73. Watts, DJ, Strogatz, SH: Collective dynamics of 'small-world' networks. Nature. **393**(6684), 440–442 (1998)
74. Federal Reserve System: The Federal Reserve System purposes and functions. Ninth edn. Board of Governors of the Federal Reserve System, Washington DC (2005)
75. Krackhardt, D: QAP partialling as a test of spuriousness. Soc. Netw. **9**, 171–186 (1987)
76. Dekker, D, Krackhardt, D, Snijders, T: Multicollinearity robust QAP for multiple regression. In: 1st Annual Conference of the North American Association for Computational Social and Organizational Science, pp. 22–25. NAACSOS, Pittsburgh, PA, (2003)
77. Dekker, D, Krackhardt, D, Snijders, TA: Sensitivity of MRQAP tests to collinearity and autocorrelation conditions. Psychometrika. **72**(4), 563–581 (2007)
78. Kilduff, M, Krackhardt, D: Bringing the individual back in: a structural analysis of the internal market for reputation in organizations. Acad. Manage. J. **37**(1), 87–108 (1994)
79. Corman, SR, Kuhn, T, McPhee, RD, Dooley, KJ: Studying complex discursive systems. Hum. Commun. Res. **28**(2), 157–206 (2002)
80. Broekel, T, Boschma, R: Knowledge networks in the Dutch aviation industry: the proximity paradox. J. Econ. Geogr. **12**(2), 409–433 (2012). doi:10.1093/jeg/lbr010
81. Xiang, Z, Gretzel, U, Fesenmaier, DR: Semantic representation of tourism on the internet. J. Travel Res. **47**(4), 440–453 (2009)
82. Ehrmann, M, Fratzscher, M: Communication by central bank committee members: different strategies, same effectiveness? J. Money, Credit and Bank. **39**(2), 509–541 (2007)
83. Friedman, BM: The use and meaning of words in central banking: inflation targeting, credibility, and transparency. NBER Working Paper, No 8972 (2002). doi:10.3386/w8972
84. Sturm, J, Haan, JD: Does central bank communication really lead to better forecasts of policy decisions? New evidence based on a Taylor rule model for the ECB. Review of World Economics. **147**(1), 41–58 (2011)
85. Valente, TW, Coronges, K, Lakon, C, Costenbader, E: How correlated are network centrality measure? Connections. **28**(1), 16–26 (2008)