A Webometric analysis of the global banking industry

Esteban Romero-Frías
erf@ugr.es

Department of Accounting and Finance, University of Granada, Granada, Spain

Liwen Vaughan
lvaughan@uwo.ca

Faculty of Information and Media Studies, University of Western Ontario, London, Ontario, Canada

Abstract

Previous research applied Webometric techniques to the study of companies and found that links to commercial Websites contain useful business information. Two main techniques have been applied successfully: inlink analysis and co-link analysis. The former showed that hyperlink counts are significantly correlated with financial variables. The latter used the number of links to a pair of business Websites (co-links) to map business competitive positions. This paper combines both techniques to offer a complete analysis of the international banking industry through the Web hyperlink data retrieved through a major search engine. Top 50 international banks from 15 different countries are candidates of the study. Correlations between link data and some financial variables (e.g. total assets, net income, total revenue) are statistically significant. Also, the maps based on co-link data indicate that geographical and linguistic factors determine competitive clusters in the international banking industry.

Keywords Competitive intelligence, Web data mining, co-link analysis, Webometrics, financial position, financial performance, banking industry.

Paper type Workshop paper

1.- Introduction

Since its creation about 20 years ago, the World Wide Web (the Web) has revolutionised the Internet, facilitating the access of all types of potential users and the creation of all type of contents. Today, the available collection of information makes possible to define the Web as an enormous, unstructured and heterogeneous database that, despite its appearance, is not randomly built and can be exploited from different perspectives. Web data mining techniques have been applied to three types of data: Web content, Web structure and Web usage. In this context, a new discipline, Webometrics (Almind and Ingwersen, 1997; Björneborn and Ingwersen, 2004), has emerged from the Information Science field to study quantitatively the Web by applying Bibliometric techniques. Hyperlink is a key raw material in Webometric studies. It is defined as a reference or navigational element in a document to another section of the same document or to another document that may be on a different domain or part of the same domain. Hyperlinks constitute the structure of the Web and can be regarded as an endorsement of a target page when created for meaningful purposes.

Exploitation of hyperlinks is well illustrated by the functioning of commercial search engines. In this sense, Google's market dominance derives from its technological lead, part of which was established by the *Pagerank* system which is based mainly on two principles (Brin and Page, 1998):

- Webpages that have more links pointing to them are considered to be more relevant.
- Not all the hyperlinks have the same value. Those links established by relevant webpages are more valuable than the others from not so well known webpages.

However, the idea that a link pointing to a Webpage means a vote to that Webpage or document is not new. Garfield (1979) proposed this reasoning as the basis to rank scientific production. Webometrics takes up this tradition and adapts a set of methods, originally designed for Bibliometric analysis of scientific journal article citation patterns, to the Web (Thelwall, Vaughan and Björneborn, 2005).

Inlink and co-inlink are basic Webometric concepts to understand the nature of Web hyperlinks. An inlink, also called back link, is a link pointing to a Webpage, e.g. page X has an inlink coming from page Z. If page X and page Y both have inlinks from page Z, then page X and page Y are co-inlinked (Björneborn and Ingwersen, 2004). Co-inlink analysis is referred simply as co-link analysis later in this paper.

The application of Webometric techniques to commercial websites is in a developing phase. Competitive Intelligence (CI) is one of the areas where this research have been more successful and promising (Tan *et al*., 2002; Reid, 2003). According to Kahaner (1996), CI consists of a systematic plan to obtain and analyse information about competitors and general trends in the industry. The abundance of digital information available on the Internet has created new opportunities and challenges for companies, therefore they need to monitor changes around them in order to compete in better conditions.

The purpose of this research is to analyse the international banking industry by using two different Webometric techniques, inlink analysis and co-link analysis. These two techniques have been proven to be effective in gathering business information in addition to traditional sources. Data gathered through

links pointing to Webpages of banks are expected to be a useful additional source of information to generate business knowledge, especially in order to study business competition within the industry. The banking industry has been selected due to its economic impact and degree of internationalization, especially in the context of the recent financial and economic crisis. Moreover, few Webometric studies have focused on industries other than the Information Technology and Telecommunication industry.

Link impact analysis (Thelwall, 2004) is based on the number of web pages that link to a collection of pages or sites in order to assess their impact. Previous research by Vaughan and colleagues used this technique to explore quantitative relationships between inlink counts to commercial Websites and business performance variables. If a significant correlation exists between these variables, then inlink counts could be used as an indicator of business performance. Vaughan and Wu (2004) tested the hypothesis in two groups of Chinese companies: one homogeneous group (in terms of industry) made of China's top 100 Information Technology (IT), and one heterogeneous group made of top 100 privately owned companies. Spearman correlation tests showed significant relationships between inlink counts and three accounting variables: gross revenue, profit, and research and development expenses. However, these relationships are only significant when homogeneous group of companies are being compared. Vaughan (2004a; 2004b) reinforced this evidence by studying IT industry in China, USA and Canada. Romero Frías, Vaughan and Rodríguez Ariza (2009) extended the research to other industries and found evidence of significant correlations between inlink counts and financial variables (total assets, revenue and net income) in the US banking industry.

Colink research is based upon the number of webpages that link at the same time two webpages or sites belonging to the set of entities under study. Co-links are analogous to the bibliometric concept of co-citation (Small, 1973). The number of co-links to the Websites of a pair of companies is a measure of the similarity between two companies. This means that the more co-links the two companies have, the more related they are in the views of the sites that link to them. When applied to a set of companies belonging to the same industry, similar companies are understood as competing businesses. Although links among academic websites have proved to be a relevant measure of similarity (Thelwall and Wilkinson, 2004), competitors link each other very rarely in order to avoid diverting Web traffic to rival companies (Shaw, 2001; Vaughan, Gao and Kipp, 2006). This fact leads to the use of co-links to research competitive positions in an industry.

Vaughan and You (2006) applied co-link analysis to map successfully the business competitive positions of 32 telecommunication companies. A more recent paper (Vaughan and You, 2008) proposed a method that combines page content mining (keyword) and Web structure mining (co-link data) to achieve a more detailed picture of a particular sector (WiMAX) within the telecommunication industry. The methodology developed in these papers has also been tested and verified in other countries and other industries, e.g. China’s chemical industry and electronics industry (Vaughan, Tang and Du, 2009). Recently, Romero-Frías and Vaughan (2009) extended the use of co-link analysis into the banking industry in the US in order to test the feasibility of combining page content with co-link data to monitor financial crisis.

This study analyses the international banking industry through a combination of Webometric techniques. First of all, we wanted to find out if there is any correlation between the number of Web hyperlink that a bank attracts and the bank’s financial variables. Secondly, we analyzed competitive positions of the banks in the study and compared the global market situation on two occasions , December 2008 and June 2009.

Finally, we tried to study to what extent financial crisis could have affected the changes in the Web data analysed. The paper is exploratory in nature and tries to present an overall picture of a global industry by using Web hyperlink data.

2.- Methodology

2.1.- Selection of companies to study

Top 50 banks of the world in terms of total assets (consolidated figures) were selected for the study. The list of companies was gathered on December 4, 2008 from http://www.bankersalmanac.com/addcon/infobank/wldrank.aspx. These banks come from 15 different countries as shown in Table 1. The complete list of all banks, including URLs, labels used in the analysis and inlink counts, is shown in Appendix 1.

Table 1. Country Distribution of the Banks

Country	Australia	Belgium	Canada	China	Denmark	France	Germany	Italy
Number of banks	1	2	2	4	1	6	8	2
Country	Japan	Netherlands	Spain	Sweden	Switzerland	UK	USA	
Number of banks	5	3	2	1	2	6	5	

The Website address of each of these banks was collected using Google and then manually checked to ensure its correctness. The vast majority of companies in the study have only one URL for their Websites. For the few that have alternative URLs in the form of alias or redirect, we checked each URL to find out which one has more inlinks and used that one for collecting inlink data. We considered including both URLs in data collection. However, Yahoo!, the search engine used for data collection, cannot handle the complex query syntax for collecting co-link data using two URLs.

2.2. Collecting financial data

As stated in the Introduction, we wanted to find out if there is any correlation between the number of Web hyperlinks that a bank attracted and the bank's financial variables. Thus we need to collect financial data of these banks. We used Mergent (http://www.mergent.com) as it is a reliable source of business information on global publicly listed companies. At the time we started the study, December 2008, the latest financial data available were for the year 2007. We considered the following three groups of financial variables:

- Financial position variables: Total Assets and Total Liabilities.
- Financial performance variables, in absolute terms: Total Revenue, Net Income and Earnings before Tax.
- Financial performance variables, in relative terms: Return on Assets (ROA).

Regarding the financial variables in the study, we were able to gather data for 39 banks out of 50. However, for the variable Earnings before Tax only 37 banks were available.

2.3. Collecting Web link data

Of the three major search engines, Google, Yahoo! and MSN Live Search, only Yahoo! could be used for data collection for the study. Google's inlink search only returns a sample of all inlinks that the Google database records (Google, 2009). Another problem is that Google cannot filter out internal inlinks (inlinks originated from within the Website itself such as "back to home" type of links) as the query term 'link" cannot be combined with any other query terms (Google, 2006). In other words, it cannot report the external inlink counts that the study needed. MSN Live Search used to have inlink search functions but the service was turned off around March 2007 (Live Search, 2007). At the time of data collection, December 2008 and June 2009, Yahoo! is the only option for collecting inlink data as required by the study.

Because search engines of different countries may have databases that favour Websites of the host countries (Vaughan and Thelwall, 2004), we considered using the versions of Yahoo! of the country to which the banks belong. This is feasible for countries such as China as the Chinese version of Yahoo! (www.yahoo.cn) has a database that is different from the global Yahoo! (www.yahoo.com). However, for other countries such as Spain and France, this is not the case. Tests of these versions of Yahoo! showed that they returned the same inlink search results as that from the global version of Yahoo!. This means that the Spanish and French version of Yahoo! just have a different interface but the same underlying database as the global Yahoo!. So the global version of Yahoo! (www.yahoo.com) was used for all data collection.

Yahoo! has two inlink search query terms, link and linkdomain. The "link" query term finds links to a particular page (e.g. link:http://www.abc.com finds links to the homepage of www.abc.com) while the linkdomain query term retrieves all links that point to all pages of a particular Website or domain including the homepage. We used the linkdomain query term for data collection because all links, not just links to homepage, are of relevance to the study. The query syntax for the data collection is illustrated in Table 2 using the hypothetical URLs of www.abc.com and www.xyz.com. We truncated the www portion of the URLs in the queries to capture all links to all subdomains such as mail.abc.com. The "-site:abc.com" part of the query is to filter out internal links coming from within the domain of abc.com itself. Since co-links involve a pair of Websites, the co-link data are collected in the form of a matrix with row x and column y of the matrix representing the number of co-links between URL x and URL y.

Table 2. Illustration of Yahoo! Queries for Data Collection

Types of links searched for	Query
Inlinks to www.abc.com	linkdomain:abc.com –site:abc.com
Co-links between www.abc.com and www.xyz.com	(linkdomain:abc.com –site:abc.com) (linkdomain:xyz.com –site:xyz.com)

We collected first round of Web hyperlink data on December 4, 2008. At that time, the financial crisis of the world banking industry and the worldwide economic recession was at its peak. We analyzed the data and found some interesting findings as reported in the *Results* section. We then collected the same type of Web hyperlink data again in June 2009. We wanted to find out if the change of the financial situation would be reflected through the Web hyperlink data.

2.4. Methods of data analysis

Co-link matrix of each round of data was analyzed using multidimensional scaling (MDS) to generate a MDS map. MDS uses a heuristic method to place banks with higher co-link counts closer in the resulting MDS map. The logic of our analysis is that on a macro level, co-links are created for a reason and thus have a pattern. Banks that have similar or related business/services are more likely to be co-linked. In other words, the number of co-links between a pair of banks could potentially be a measure of their similarity or relatedness. The more co-links, the more similar or related would be the business/services of the two banks. Since similar or related banks will be placed closer in the MDS map and banks with similar or related business/services are competitors (two banks offer totally different services are not competing with each other), the MDS map will group competing banks together. We hoped to use MDS the map to see clusters of banks that reveal their market positions. We also hoped to examine of the effectiveness of our methods by comparing the MDS maps from different time periods. If the MDS map would show business relationships among banks, maps from different time periods could potentially reflect changes in business situations.

The raw co-link count collected from Yahoo! needed to be normalized to obtain a relative measure of the relatedness of companies. This is necessary because a co-link count of ten is large if the two companies in question each had very few inlinks, e.g. 15. On the other hand, if each company received thousands of inlinks, then the 10 co-links that they shared represented a very small portion. The normalization was done by applying the Jaccard index in the following way.

$$\text{Normalized co-link count } = n(A \cap B) / n(A \cup B)$$

Where A is the set of the Web pages that linked to Website X

B is the set of the Web pages that linked to Website Y

$n(A \cap B)$ is the number of pages that linked to both Website X and Website Y, i.e. the raw co-link count

$n(A \cup B)$ is the number of pages that linked to either Website X or Website Y.

The normalized co-link matrices were feed into SPSS for MDS analysis. The stress values of MDS analysis are 0.075 for the first round of data and 0.073 for the second round of data. These stress values are all fairly low, which suggests a good fit between the data and the MDS map positions.

3.- Results

3.1.- Correlation between Web inlink and financial data

Before carrying out the correlation tests, we first performed descriptive statistical analysis for both rounds of inlink data. For the December 2008 data, the mean and the median are 133,396 and 52,100 respectively. For the June 2009 data, they are 131,564 and 42,700 respectively. The numbers did not

change much over the six month of time period, suggesting that inlink counts are a fairly stable over time. However, if we analyze the changes at individual bank level, we find that inlink counts for some companies have changed significantly, as shown in Table 3.

Table 3. Top changes in inlink counts from December 2008 to June 2009

	Top 5 increases in inlinks		Top 5 decreases in inlinks	
1	China Construction Bank Corporation	77.30%	The Norinchukin Bank	-54.06%
2	DZ Bank AG	48.19%	Wachovia Bank NA	-41.33%
3	Kreditanstalt für Wiederaufbau (KfW)	41.53%	Calyon	-39.54%
4	Agricultural Bank of China	41.45%	Fortis Bank SA/NV	-39.12%
5	Deutsche Bank AG	33.28%	Dresdner Bank Group	-39.02%

German banks are some of the banks that increased more the number of inlinks over the last six months. This can be explained by some economic facts. For instance, KfW (41.53% more inlinks) is back to earning profits in the first quarter of 2009 after two years of heavy losses resulting from the crisis of financial markets. Also, the interim accounts of the KfW Group at 31 March 2009 closed with a consolidated profit of EUR 80 million. Finally, Deutsche Bank (33.28%) reported that net income for the first quarter 2009 was EUR 1.2 billion, compared to a net loss of EUR 141 million in the same period of 2008. Among the top five banks with increased inlinks, two are from China. It is known that the Chinese banks overall weathered the financial crisis better because of tighter government regulations (Shaw and Parussini, 2009).

On the other side, banks, such as Wachovia, Fortis and Dresdner, reduced the number of inlinks in the same period. This situation could be interpreted as investors and other stakeholders losing attention on these companies. Again some economic information can help to understand the reasons why. Wachovia Corporation (-41.33%) was purchased by Wells Fargo on December 2008, and consequently it ceased to be an independent corporation on that date. The Wachovia brand is expected to be absorbed into the Wells Fargo brand over the next three years. Fortis (-39.12%) underwent serious financial problems over the last 2 years and finally most of the company was sold in parts in 2008. Finally, Commerzbank announced that it would acquire Dresdner Bank (-39.02%) on August 31, 2008. The transaction is expected to be completed by the end of 2009.

The frequency distribution of the two sets of inlink counts (December 2008 and June 2009) are very skewed, so the Spearman correlation test rather than the Pearson correlation test was used. The correlation coefficients between the inlink counts and the financial performance data are shown in Table 4. All correlation coefficients are statistically significant except for the variable ROA.

Table 4. Correlation between inlink data and financial data

	Total Assets	Total Liabilities	Total Revenue	Net Income	Earnings before Tax	ROA
Dec 2008 inlink data	0.44**	0.43**	0.44**	0.57**	0.57**	0.02
June 2009 inlink data	0.43**	0.42**	0.40*	0.55**	0.57**	0.06

**. Correlation is significant at the 0.01 level (2-tailed). *. Correlation is significant at the 0.05 level (2-tailed).

Findings show that inlink counts could be use as an indicator of the banks' financial position and financial performance measures in absolute terms. However, there is no significant relationship between inlink count and ROA, a relative financial performance measure. This makes sense since the number of inlinks retrieved from the commercial search engine includes all the links pointing to a particular bank Website since its creation. This means that the inlink count is accumulative in nature. The same could be said about the financial variables in absolute terms, especially for the total assets and liabilities variables. It is expected that a company becomes bigger with the time and therefore assets and liabilities increase. Performance variables (e.g. total revenue and net income) are calculated for a particular period of time. However, it is also expected that revenue and net income will increase year after year, e.g. total revenue is expected to be higher in the fifth year of a company than in its first year. By contrast ROA is not dependent of the company's size and therefore it is logical that no significant correlation is found. It would be interesting to analyse the changes of inlink counts year after year to calculate also a relative measure for this variable.

Findings mean that the bigger the bank's size and the results in absolute terms are, the more inlinks a bank's Website attracts . The correlation is stable over time as suggested by the very close match between the two sets of correlation coefficients. However it is slightly higher for the inlink data collected on December 2008, closer to the financial period considered in the study, 2007.

These correlation coefficients are consistent with the coefficients reported on Romero Frías, Vaughan and Rodríguez Ariza (2009) for the US banking industry, as show in Table 5.

Table 5. Correlation between inlink data and financial data for the US banking industry (Romero Frías, Vaughan and Rodríguez Ariza, 2009)

	Total Assets	Total Liabilities	Total Revenue	Net Income	ROA
January 2009 inlink data	0.74**	0.73**	0.75**	0.63**	0.13
May 2009 inlink data	0.70**	0.69**	0.71*	0.62**	0.18
**. Correlation is significant at the 0.01 level (2-tailed).					

Correlation coefficients for US banking industry are higher than for the global banking industry. This is explained by the homogeneous competitive conditions that exist in the US market in comparison to the heterogeneous markets where the banks included in this study operate. For instance, the extent to which Internet is used for commercial purposes in the different countries could explain a different degree of correlation coefficients.

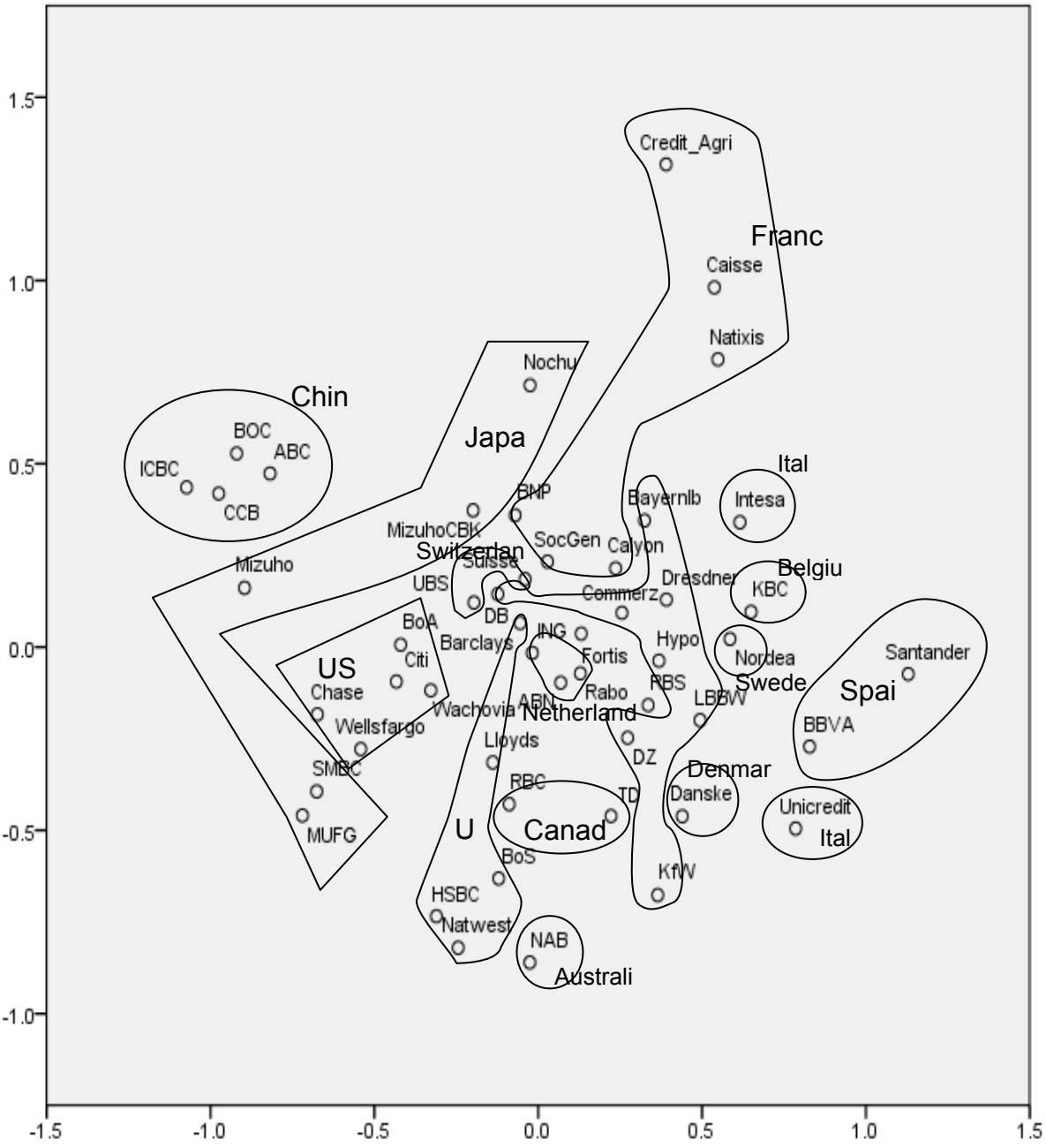
The correlation does not mean causation though. The large number of inlinks that a bank's Website attracted does not cause the better financial performance, although a positive Web image as suggested by the larger number of inlinks may contribute positively to the bank's business. A likely explanation of the correlation is that a bank that is doing well would have better financial data and also being able to maintain a high profile on the Web attracting larger number of inlinks. In other words, the larger number

of inlinks may be a symptom rather than a direct cause of the good economic performance of a bank. Although we cannot establish a causal relationship, the correlation found is still very useful information. Now that we know that the two variables are correlated, we can predict one based on the other. Because the Web inlink data are publically accessible and can be collected easily, we could find trends of a bank's financial performance in the long run based on inlink data. The fact that we were not able to locate financial data for some banks in the study underscores the usefulness of this approach. For non listed companies, financial data are most likely unavailable. The correlation that we found is potentially more useful in situation of scarce or no information available. Another use of the correlation found in the study is to identify companies whose Web presence is not on par with their financial performance. The correlation could be used to develop a regression equation and we could use the equation to predict what a company's Web position should be based on their financial data. Comparing the predicted Web position with the real one would reveal those that are underperforming on the Web. Further analysis could then be done to find out why and how to improve.

3.2. Co-link analysis

MDS maps (see Figure 1 and 2) based on the number of co-links between banks show the relative competitive positions of the companies in the study. An MDS map depicting the positions of companies in the global market in December 2008 is shown in Figure 1. This MDS map presents a clear pattern that reflects national, language and regional clusters. Three main areas are identified: Asian banks, English language banks and European banks (excluding UK). Chinese and Japanese banks are clustered together in one side of the map. Chinese banks are located at the outskirt of the map isolated from other companies. This indicates that they compete mainly at local level. Also Chinese banks receive many inlinks but relatively few of them are co-links with other companies. Japanese banks are in an intermediate position with MUFG and SMBC close to US banks, Mizuho to Chinese banks and Nochu and MizuhoCBK to European entities. Banks from English speaking countries (US, UK, Canada and Australia) are clustered together, although national groups are also clearly identified within the cluster. European banks (from Belgium, Denmark, France, Germany, Italy, Netherlands, Spain, Sweden and Switzerland) occupy the rest of the map. The two Swiss banks are in a central position among all banks studied, mirroring the central and competitive position of the Swiss banks. Dutch banks and some companies that are not cluster with their national banks (Fortis and Royal Bank of Scotland) are also in a central position. French banks are located in the same area: three of them -BNP, Calyon and Société Générale- are very close to other European banks, whereas others -Caisse d'Epargne, Crédit Agricole and Natixis- are at the outskirt of the map. Like the French banks, the German ones are together but spread over the map.

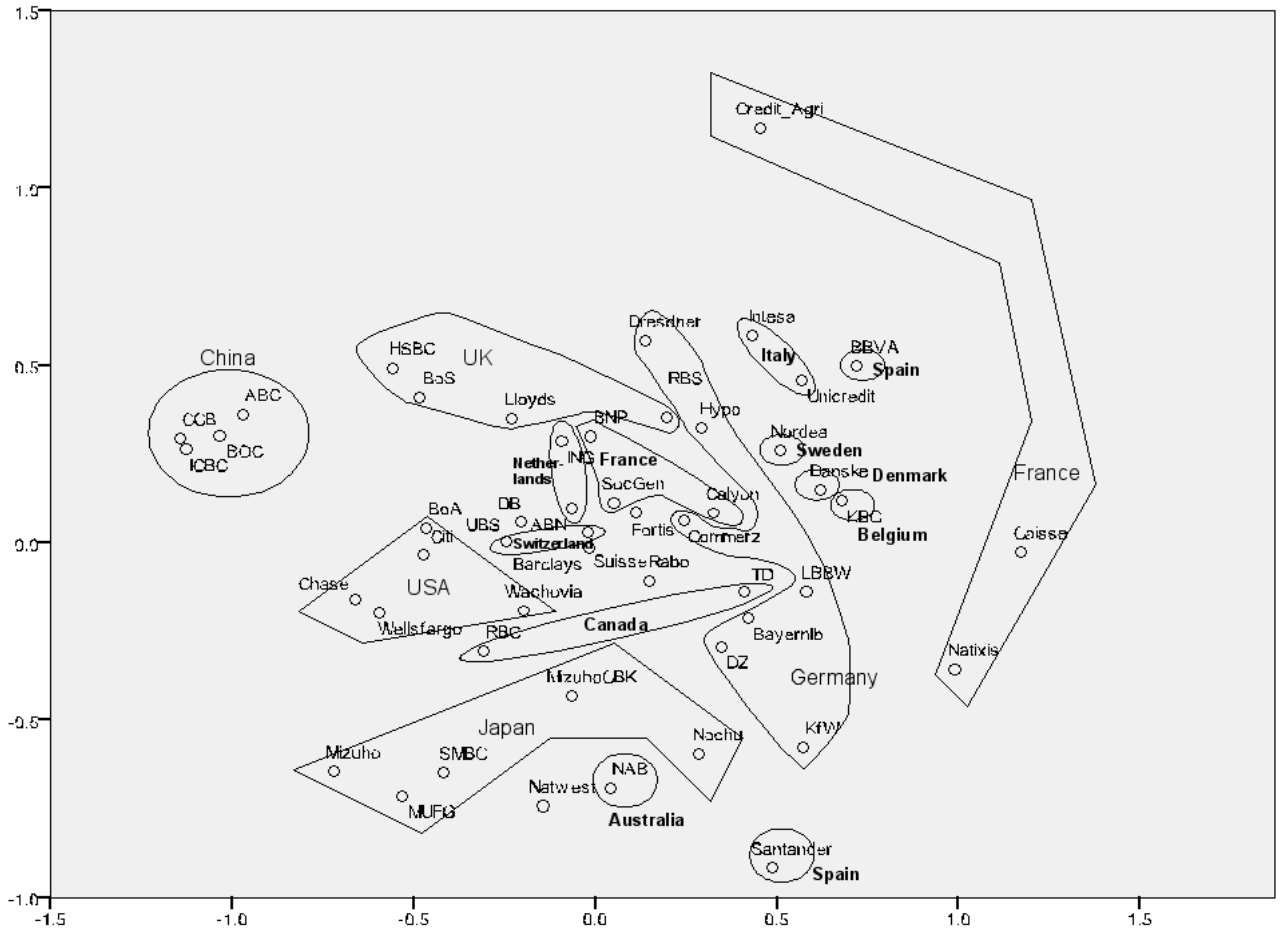
Figure 1. MDS map for banking industry (December 2008)



The MDS map reflecting the banking industry at June 2009 is shown in Figure 2. To be consistent with the previous map, clusters based on the origin country of the companies are shown. Despite the fact that this factor remains as the most consistent factor to cluster the banks, there are some significant changes that are likely to be related to the evolution in the economic and financial crisis over the last six months. It is worth noting that the Chinese and Japanese banks moved apart and the US banks are placed between them. Further monitoring and analysis is need to find out whether this reflects the changing perception the world may have on the Chinese banks as the result of the world financial crisis. British banks are closer to

Chinese banks than six months before. Swiss banks and major Europeanbanks (Deutsche Bank and Barclays) maintain their central positions. This reflects the strong and stable positions these banks have in the industry. French banks are now divided in two groups. In line with the previous analysis, three of the French banks are located in the centre and other three at the outskirt of the map. German banks do not show any significant change. It is worth to note that in December 2008 the two Italian banks were away in the map and the Spanish ones together. In June 2009 the situation is the opposite.

Figure 2. MDS map for banking industry (June 2009)



4. Conclusions and future research

The original purpose of this study was to use a combination of inlink and co-link analysis to provide a general analysis of the global banking industry. The significant correlations found between inlink counts and financial variables show that Web structure data can provide useful business information. Findings are consistent with previous research in the same industry (Romero Frías, Vaughan and Rodríguez Ariza, 2009). Despite the internationalization process of financial activities, MDS analysis based on co-links indicates that the global banking industry is partially organised in regional markets despite the existence

of main global players. Some of them seem to be more isolated such as the Chinese market. As expected, Swiss banks are located at the center of the map, together with major banks (e.g. Barclays or Deutsche Bank) from different European countries. The comparison between the December 2008 map and the June 2009 map suggests the existence of some relevant differences that need to be analysed at the light of the changes undergone in particular countries due to the world economic and financial crisis. To improve the co-link analysis, we plan to perform some interviews with financial analysts who are experts in the industry.

There could be some limitations in the inlink and co-link analysis due to language differences, especially in markets where Yahoo! maintains a different database (e.g. China). Also, the interpretation of the results could differ due to factors such as alternative criteria to group the banks, or specific knowledge of the researcher about a particular economic context. Also the number of banks in each country is not homogeneous and therefore this could affect the results as the main variable used in the analysis is the bank's country origin. It is also worth to note that the Web policy of the banks is a significant issue. A bank could operate in multiple countries by using different domains or under the same domain or brand, conditioning the analysis undergone in this paper.

Future research will focus on monitoring the changes of the MDS maps by collecting and analyzing data over time. Also, although genuine Webometric techniques are quantitative, research could be complemented by a content analysis of the inlinking and co-linking Webpages. Finally, it is necessary to develop a systematic methodology to analyse industries or companies using the different Webometric techniques available.

References

Almind, T. C. and Ingwersen, P. (1997). Informetric analyses on the World Wide Web Methodological approaches to 'webometrics', *Journal of Documentation*, 53(4), 404-426.

Björneborn, L. and Ingwersen, P. (2004). Toward a basic framework for webometrics, *Journal of the American Society for Information Science and Technology*, 55(14), 1216–1227.

Brin, S. and Page, L. (1998). The anatomy of a large-scale hypertextual Web search engine, *Computer Networks and ISDN Systems*, 30, 1-7.

Garfield, E. (1979). *Citation indexing: Its theory and applications in science, technology and the humanities*. New York: Wiley.

Google (2006). *Google SOAP Search API Reference*. Retrieved June 3, 2009, from http://www.google.com/apis/reference.html#2_2

Google (2009). *Links to your site*. Retrieved June 3, 2009, from http://www.google.com/support/webmasters/bin/answer.py?hl=en&answer=55281

Kahaner, L. (1996). *Competitive Intelligence – How to Gather, Analyze, and Use Information to Move Your Business to the Top*, 7th ed., New York: Touchstone.

Live Search (2007). *We are flattered, but...* Retrieved June 3, 2009, from http://www.bing.com/community/blogs/search/archive/2007/03/28/we-are-flattered-but.aspx

Reid, E. (2003). Using Web link analysis to detect and analyze hidden Web communities. In D. Vriens (ed.), *Information and Communications Technology for Competitive Intelligence*, Ohio: Ideal Group Inc, 57-84.

Romero Frías, E., Vaughan, L. and Rodríguez Ariza, L. (2009). El recuento de enlaces a sitios Web comerciales como indicador de las variables de desempeño y posición financiera de la empresa: estudio empírico de diversos sectores empresariales en Estados Unidos, *XV Congreso de la Asociación Española de Contabilidad y Administración de Empresas*, Valladolid, September 23-25, 2009.

Shaw, D. (2001). Playing the links: interactivity and stickiness in .com and "not.com" websites. *First Monday*, 6(3). Retrieved May 10, 2009, from http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/837/746

Shaw, J. and Parussini, G. (2009). China bank regulator: Global regulatory cooperation very fragile. Retrieved Oct. 15, 2009 from http://www.nasdaq.com/aspx/stock-market-news-story.aspx?storyid=200910051100dowjonesdjonline000261&title=china-bank-regulatorglobal-regulatory-cooperation-very-fragile.

Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for Information Science*, July-August, 265–269.

Tan, B., Foo, S. and Hui, S. C. (2002). Web information monitoring for competitive intelligence. *Cybernetics and Systems*, 33(3), 225-251.

Thelwall, M. (2004). *Link Analysis: An Information Science Approach.* San Diego: Academic Press.

Thelwall, M. and Wilkinson, D. (2004). Finding similar academic Web sites with links, bibliometric couplings and colinks. *Information Processing and Management*, 40, 94–101.

Thelwall, M., Vaughan, L. and Björneborn, L. (2005). Webometrics. In B. Cronin (ed.), *Annual review of information science and technology*, Medford, NJ: Information today, 39, 81-135.

Vaughan, L. (2004a). Exploring website features for business information. *Scientometrics*, 61(3), 467-477.

Vaughan, L. (2004b). Web hyperlinks reflect business performance—A study of US and Chinese IT companies. *Canadian Journal of Information and Library Science*, 28(1), 17–31.

Vaughan, L. and Thelwall, M. (2004). Search engine coverage bias: Evidence and possible causes. *Information Processing & Management*, 40(4), 693-707.

Vaughan, L. and You, J. (2006). Comparing business competition positions based on Web co-link data—The global market vs. the Chinese market. *Scientometrics*, 68(3), 611–628.

Vaughan, L. and You, J. (2008). Content assisted web co-link analysis for competitive intelligence. *Scientometrics*, 77(3), 433-444.

Vaughan, L., and Wu, G. Z. (2004). Links to commercial websites as a source of business information. *Scientometrics*, 60(3), 487–496.

Vaughan, L., Gao, Y. and Kipp, M. (2006). Why are hyperlinks to business Websites created? A content analysis. *Scientometrics*, 67(2), 291–300.

Vaughan, L., Tang, J. and Du, J. (2009). Examining the robustness of Web co-link analysis. To appear in *Online Information Review*, 33(5).

Appendix 1. Banks included in the study

Labels	Company	Main URL	Country	Dec 2008 inlink data	June 2009 inlink data
RBS	The Royal Bank of Scotland Group plc	http://www.rbs.com	UK	26,700	35,200
DB	Deutsche Bank AG	http://www.db.com	Germany	66,100	88,100
BNP	BNP Paribas SA	http://www.bnpparibas.com	France	143,000	112,000
Barclays	Barclays PLC	http://www.barclays.com	UK	42,900	28,200
Credit-Agri	Crédit Agricole SA	http://www.credit-agricole.com	France	28,700	30,900
UBS	UBS AG	http://www.ubs.com	Switzerland	147,000	148,000
SocGen	Société Générale	http://www.socgen.com	France	54,300	42,000
ABN	ABN AMRO Holding NV	http://www.abnamro.com	Netherlands	52,100	42,700
Unicredit	UniCredit SpA	http://www.unicreditgroup.eu	Italy	79,900	50,000
ING	ING Bank NV	http://www.ing.com	Netherlands	58,800	64,900
MUFG	The Bank of Tokyo-Mitsubishi UFJ Ltd	http://www.mufg.jp	Japan	364,000	454,000
Santander	Banco Santander SA	http://www.santander.com	Spain	395,000	364,000
Chase	JPMorgan Chase Bank National Association	http://www.chase.com	USA	223,000	144,000
BoA	Bank of America NA	http://www.bankofamerica.com	USA	339,000	304,000
Citi	Citibank NA	http://www.citibank.com	USA	140,000	98,700
Suisse	Credit Suisse Group	http://www.credit-suisse.com	Switzerland	61,500	70,100
Fortis	Fortis Bank SA/NV	http://www.fortis.com	Belgium	34,000	20,700
ICBC	Industrial & Commercial Bank of China Limited	http://www.icbc.com.cn	China	985,000	937,000
CCB	China Construction Bank Corporation	http://www.ccb.com	China	304,000	539,000
BoS	Bank of Scotland plc	http://www.bankofscotland.co.uk	UK	13,700	9,390
HSBC	HSBC Bank plc	http://www.hsbc.co.uk	UK	85,100	65,700
Intesa	Intesa Sanpaolo SpA	http://www.intesasanpaolo.com	Italy	47,800	46,700
SMBC	Sumitomo Mitsui Banking Corp.	http://www.smbc.co.jp	Japan	105,000	103,000
Commerz	Commerzbank AG	http://www.commerzbank.com	Germany	10,600	9,100
Calyon	Calyon	http://www.calyon.com	France	10,900	6,590

Rabo	Rabobank Nederland	http://www.rabobank.com	Netherlands	11,100	11,300
Dresdner	Dresdner Bank Group	http://www.dresdner-bank.com	Germany	5,280	3,220
Caisse	Caisse Nationale des Caisses d'Epargne et de Prévoyance	http://www.caisse-epargne.com	France	28,800	31,700
Lloyds	Lloyds TSB Group plc	http://www.lloydstsb.com	UK	40,400	28,900
ABC	Agricultural Bank of China	http://www.abchina.com	China	193,000	273,000
BOC	Bank of China Limited	http://www.boc.cn	China	399,000	485,000
Danske	Danske Bank A/S	http://www.danskebank.com	Denmark	7,990	7,820
Wachovia	Wachovia Bank NA	http://www.wachovia.com	USA	143,000	83,900
RBC	Royal Bank of Canada	http://www.royalbank.com	Canada	40,600	40,500
Hypo	Bayerische Hypo-und Vereinsbank AG	http://www.hypovereinsbank.de	Germany	14,000	16,300
Natixis	Natixis	http://www.natixis.fr	France	7,690	4,720
Nochu	The Norinchukin Bank	http://www.nochubank.or.jp	Japan	5,790	2,660
DZ	DZ Bank AG	http://www.dzbank.de	Germany	9,650	14,300
Natwest	National Westminster Bank Plc	http://www.natwest.com	UK	21,800	18,300
Nordea	Nordea Group	http://www.nordea.com	Sweden	41,200	35,900
Mizuho	Mizuho Bank Ltd	http://www.mizuhobank.co.jp	Japan	124,000	113,000
LBBW	Landesbank Baden-Württemberg	http://www.lbbw.de	Germany	6,190	7,070
MizuhoCBK	Mizuho Corporate Bank Ltd	http://www.mizuhocbk.co.jp	Japan	5,710	3,570
KfW	Kreditanstalt für Wiederaufbau	http://www.kfw.de	Germany	44,300	62,700
BBVA	Banco Bilbao Vizcaya Argentaria SA	http://www.bbva.com	Spain	43,100	53,200
NAB	National Australia Bank Ltd	http://www.nab.com.au	Australia	19,800	21,200
Wellsfargo	Wells Fargo Bank NA	http://www.wellsfargo.com	USA	346,000	291,000
Bayernlb	Bayerische Landesbank	http://www.bayernlb.de	Germany	5,810	4,700
KBC	KBC Bank NV	http://www.kbc.com	Belgium	17,000	20,300
TD	The Toronto-Dominion Bank	http://www.td.com	Canada	37,300	39,800