Web hyperlink patterns and the financial variables of the global banking industry

Liwen Vaughan [1]

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

Esteban Romero-Frías

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

Abstract

The goal of the study is to explore business and social information contained in Web data and to examine the effectiveness of Webometrics methods in obtaining this type of information. Specifically, this study combines the inlink and co-link analysis techniques to make a complete examination of the international banking industry. Results from the analysis were compared with the banks' real financial situation to determine the validity and reliability of the link analysis methods. The top 50 international banks from 15 different countries were candidates for the study. Web hyperlink data were collected for two time periods six months apart. Financial data from two different years were collected to find out which year correlated better with inlink data. Statistically significant correlations were found between inlink data and several financial variables. A comparison between Asian banks and other banks showed that the former attracted significantly more inlinks. The multidimensional scaling (MDS) maps generated from co-link data suggest that geographic and linguistic factors determine competitive clusters in the international banking industry. A comparison of the MDS maps from the two different time periods revealed important business information, notably that the Chinese banks moved closer to the major banks from the U.S. and U.K.. This is in line with the current state of Chinese banks, which emerged from the financial crisis of 2008 in a stronger position than they had held going into it.

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

---

[1]   *Correspondence to: Liwen Vaughan, Faculty of Information and Media Studies, University of Western Ontario, London, Ontario, Canada, N6A 5B7. Email: lvaughan@uwo.ca*

1. Introduction

Much research has been carried out into the application of Webometric techniques to commercial Websites. Competitive Intelligence (CI) is one of the areas in which this type of research has been successful and promising [1, 2]. According to Kahaner [3], 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 be more competitive in the marketplace. The purpose of this research is to analyse the international banking industry by using two different Webometric techniques, inlink analysis and co-link analysis. Inlink and co-inlink are basic Webometric concepts that enable us to understand the nature of Web hyperlinks. An inlink, also called back link, is a hyperlink embedded on one page that points to another page. For example, if there is a hyperlink on page Z that points to page X, then page X has an inlink. If page X and page Y both have inlinks from page Z, then page X and page Y are co-inlinked [4]. Co-inlink analysis is hereafter referred simply as co-link analysis. These two Webometric techniques have been proven effective in gathering business information in addition to traditional sources. Data gathered through links pointing to the Webpages of banks are therefore expected to be a useful additional source of information from which to generate business knowledge, especially with regard to business competition. The banking industry was selected for the study because of its economic impact and degree of internationalization, especially in the context of the recent financial and economic crisis.

Inlink analysis [5] assess the impact of an organizations by analyzing the number of web pages that link to the organizations' Website. 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 [6] tested the hypothesis in two groups of Chinese companies: one homogeneous group (in terms of industry) comprising China's top 100 information technology (IT) companies, and one heterogeneous group comprising the 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. Vaughan [7, 8] reinforced this evidence by studying the IT industry in China, the U.S.A. and Canada. Romero Frías, Vaughan and Rodríguez Ariza [9] 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 U.S. banking industry.

Co-links are analogous to the bibliometric concept of co-citations [10]. The number of co-links to the Websites of a pair of companies is a measure of the similarity between the two companies. This means that the more co-links the two companies have, the more related they are perceived to be by the sites that link to them. When applied to a set of companies belonging to the same industry, similar companies are understood to be competing businesses. Although inlinks among academic websites have been proven to be a relevant measure of similarity [11], in business, competitors very rarely link to each other in order to avoid diverting Web traffic to rival companies [12, 13]. This leads to the use of co-links to research competitive positions in an industry.

Vaughan and You [14] successfully applied co-link analysis to map the business competitive positions of 32 telecommunication companies. A more recent paper [15] proposed a method that combines page content (keyword) and Web structure (co-link data) to achieve a more detailed picture of competition in a particular sector (WiMAX) within the telecommunication industry. The methodology developed in these papers has also been tested and verified in other countries and industries, such as China's chemical and electronics industries [16]. Recently, Romero-Frías and Vaughan [17] extended the use of co-link analysis into the banking industry in the U.S. in order to test the feasibility of combining page content with co-link data to monitor financial crisis.

This study extends these earlier studies by applying various Webometric techniques to the international banking industries. By studying an entity that embodies various factors such as economic, culture and language, we hope to explore the effectiveness of Webometrics methods in analyzing these important social variables. The significance of the study lies not on gaining particular business information about the banking industry but on contributing to our knowledge of the Webometrics methodology and our knowledge of the Web phenomenon in general. Our specific research questions are: is there any correlation between the number of Web hyperlinks that a bank attracts and the bank's financial variables; can we use co-link analysis to map these banks' global market

positions. Both inlink and co-link data were collected in two time periods, December 2008 and June 2009, to find out if and how results changed during the recent world financial crisis. We also compared the Asian banks with other banks to determine if and how they differed in hyperlink patterns. This comparison was guided by the premise of cultural difference and the rapid development of the Asian economy.

2. Methodology

2.1. Selection of companies to study

The 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	U.K.	U.S.A.	
Number of banks	5	3	2	1	2	6	5	

The banks' Website addresses were collected using Google and then manually checked to ensure their correctness. The vast majority of companies in the study had only one URL for their Websites. For the few that have alternative URLs in the form of an alias or redirect, we checked each URL to determine which one had more inlinks and used that URL to collect inlink data. We considered including both URLs in data collection; however, at the time of the study, Yahoo! (the search engine used for data collection) couldn't handle the complex query syntax for collecting co-link data using two URLs.

2.2. Collecting financial data

The banks' financial data were collected in November 2009, using the Mergent database (http://www.mergent.com), a reliable source of business information on global publicly listed companies. The latest financial data available were for the year 2008. The same set of financial data for 2007 were also collected in order to compare the correlations to inlink counts. We considered the following three groups of financial variables:

1. Financial position variables: Total Assets.

2. Financial performance variables, in absolute terms: Total Revenue and Net Income.

3. Financial performance variables, in relative terms: Return on Assets (ROA).

Financial data for some banks were not available in the Mergent database, therefore we omitted them in the correlations tests. As shown in table 4, the number of banks in each test varied from 38 to 41. Inlink data were available for all banks in the study, therefore all 50 world banks were included for the co-link analysis.

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 [18]. In addition, Google cannot filter out internal inlinks (inlinks originating from within the Website itself, such as "back to home" links) as the query term "link" cannot be combined with any other query terms [19]. In other words, it cannot report the external inlink counts required for the study. MSN Live Search used to have inlink search functions but the service was turned off around March 2007 [20]. At the time of data collection, December 2008 and June 2009, Yahoo! was the only option for collecting the required inlink data.

Because search engines from different countries may have databases that favour Websites from the host country [21], we considered using the country-specific version of Yahoo! appropriate to each bank. This is feasible for countries such as China, as the Chinese version of Yahoo! (www.yahoo.cn) has a database that is different from that of the global Yahoo! (www.yahoo.com). However, for other countries, such as Spain and France, this is not the case. Tests of the Spanish and French versions of Yahoo! returned the same inlink search results as the global version of Yahoo!, which means that the Spanish and French Yahoo! had a different interface but the same underlying database as the global Yahoo!. Consequently, 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 (for example, 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 a homepage, are relevant to the study. The query syntax for the data collection is illustrated in Table 2 using the hypothetical URLs www.abc.com and www.xyz.com. We truncated the www portion of the URLs in the queries to capture links to all subdomains such as mail.abc.com. The "-site:abc.com" part of the query filters out internal links coming from within the domain of abc.com itself. Since co-links involve a pair of Websites, the co-link data were 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 in December 2008. At that time, the financial crisis of the world banking industry and the worldwide economic recession was at its peak. We analysed the data and discovered some interesting findings, as reported in the *Results* section. We then collected the same type of Web hyperlink data again in June 2009 when the world economy began to recover. We wanted to find out if the change in the world financial situation would be reflected in the Web hyperlink data.

2.4. Methods of data analysis

A co-link matrix of each round of data was analyzed using multidimensional scaling (MDS) to generate an MDS map. MDS uses a heuristic method to place banks with higher co-link counts closer together 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 the business/services of the two banks. Since similar or related banks

will be placed closer together in the MDS map and banks with similar or related business/services are competitors (two banks that offer totally different services are not competing with each other), the MDS map would group competing banks together. We hoped to use the MDS map to see clusters of banks that would reveal the banks' 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 showed business relationships among banks, maps from different time periods could potentially reflect changes in business situations.

The raw co-link counts collected from Yahoo! were normalized by Jaccard index to obtain a relative measure of the relatedness of the banks. The normalized co-link matrices were fed 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 both 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 time period, suggesting that inlink counts are fairly stable over time. However, if we analyse the changes at the 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 banks with increased inlink count		Top 5 banks with decreased inlink count	
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 were among those with the greatest increase in inlink counts. This could be explained by some economic facts. For instance, KfW (41.53% more inlinks) returned to profitability in the first quarter of 2009 after two years of heavy losses resulting from the financial crisis. In addition, the interim accounts of the KfW Group at March 31, 2009 closed with a consolidated profit of EUR 80 million. Finally, Deutsche Bank, the biggest German bank, had a 33.2% increase in inlink count. It reported a net income of EUR 1.2 billion for the first quarter of 2009, compared to a net loss of EUR 141 million for the same period of 2008. Also, in the first quarter of 2009, Deutsche Bank accelerated its plan for taking over the country's biggest retail bank, Deutsche Postbank [22]. Among the top five banks with increased inlinks, two are from China. It is known that the Chinese banks weathered the financial crisis better overall because of their tighter government regulations [23].

In contrast, banks such as Wachovia, Fortis and Dresdner had decreased inlink counts for the same period, which could be interpreted as investors and other stakeholders losing interest in these companies. Again some economic information can provide insight into the potential reasons. Wachovia Corporation (41.33% decrease in inlink count) was purchased by Wells Fargo in December 2008, and consequently ceased to be an independent

corporation. The Wachovia brand is expected to be absorbed into the Wells Fargo brand over the next three years. Fortis (39.12% decrease in inlink count) experienced serious financial problems over the previous two years and most of the company was sold in parts in 2008. Finally, Commerzbank announced that it would acquire Dresdner Bank (39.02% decrease in inlink count) on August 31, 2008. This EUR 9 billion takeover was finalized in the first quarter of 2009 [22]. Although this move will create a second major national banking champion together with Deutsche Bank, the decrease in the number of inlinks to Dresdner Bank's Website suggests a loss of attention in the acquired entity. As indicated on Dresdner Bank's Website, it has been a brand of Commerzbank since May 2009. As for Norinchukin Bank, its situation could be understood in light of the generally poor situation of the Japanese banking sector, especially in contrast to the strength of other Asian banks, i.e., the Chinese banks [24].

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 return on assets.

Table 4. Correlation between inlink data and financial data

Spearman's rho	Financial year	N	Inlinks of December 2008	Inlinks of June 2009
Total Assets	2008	39	.58**	.54**
	2007	39	.44**	.43**
Total Revenue	2008	38	.62**	.60**
	2007	39	.44**	.40*
Net Income	2008	39	.56**	.52**
	2007	39	.57**	.55**
Return on assets (Net)	2008	39	.23	.19
	2007	39	.02	.06
Number of employees	2008	41	.80**	.80**
	2007	39	No employee data for 2007	No employee data for 2007
**. 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 used as an indicator of the banks' financial positions and financial performance measured in absolute terms. However, there is no significant relationship between inlink count and return on assets, a relative financial performance measure. This makes sense since the number of inlinks retrieved from the commercial search engine includes all of the links pointing to a particular bank's Website since its creation. This means that the inlink count is cumulative in nature. The same could be said about the financial variables that are measured in absolute terms, especially the total assets variable. It is expected that a company becomes bigger over time and therefore its assets increase. Although performance variables such as total revenue and net income are calculated for a particular period of time, they usually increase over the years. For example, total revenue is expected to be higher in a company's fifth year than in its first year. In contrast, return on assets (ROA) is a relative measure because it is calculated as net income divide by total assets. Thus ROA is not dependent on company sizes and not surprisingly there is no significant correlation between inlink counts and

ROA. It would be interesting to analyse the changes of inlink counts over several years to calculate a relative measure for this variable.

The correlation is stable over time as suggested by the very close match between the two sets of correlation coefficients. The higher correlations are found between the December 2008 inlink data and the 2008 financial data, i.e., when the inlink data year matches that of the financial data. If we compare the correlation coefficients between inlink counts and financial data from 2007 and 2008, we observe that the relationship is stronger for the 2008 financial data. The only exception is net income, which presents a very similar correlation coefficient in both cases, although slightly higher for the 2007 financial data. This indicates that matching the observation period could be an important step in obtaining a more accurate predication of one variable based on the other. It also suggests that inlink data are somewhat time sensitive, which is a desirable feature.

These correlation coefficients are consistent with the coefficients reported in Romero Frías, Vaughan and Rodríguez Ariza [9] for the U.S. banking industry, as shown in Table 5.

Table 5. Correlation between inlink data and financial data for the U.S. banking industry [9]

	Total Assets	Total Liabilities	Total Revenue	Net Income	Return on Assets
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 is significant at the 0.05 level (2-tailed).					

Correlation coefficients for the U.S. banking industry are higher than those for the global banking industry. This is explained by the homogeneous competitive conditions that exist in the U.S. market compared with the heterogeneous markets in which the banks included in this study operate. Different countries have different economical and financial conditions, which could explain the lower correlations when banks from many countries are analysed together. The extent to which the Internet is used for commercial purposes in different countries could also be a factor.

It is important to appropriately interpret the correlations found in the study. Proving that a correlation is statistically significant does not prove causation. The large number of inlinks that a bank's Website attracts does not cause 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 be able to maintain a high profile on the Web, thus attracting a 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. Now that we know that the two variables are correlated, we can predict one based on the other. Because the Web inlink data are publicly accessible and can be collected easily, we could project a bank's long-term financial performance based on inlink data. The fact that we were not able to locate financial data for some banks underscores the usefulness of this approach. For non-listed companies, financial data are most likely unavailable. The correlation that we found is therefore potentially more useful in situations in which financial data are scarce or not available at all. 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 determine the reason for such underperformance and how the company might improve.

3.2 Inlink count comparison between Asian banks and other banks

Due to the particular features of the Asian economy and the specific positions of the Chinese and Japanese banks revealed by our co-link analysis (reported later in section 3.3), we decided to determine whether or not there

is a statistically significant difference in inlink counts between Asian banks and other banks in the study. Mann-Whitney tests were carried out for this purpose. The tests indicate that inlink counts are significantly different (p=0.042 for December 2008 data and p=0.024 for June 2009 data). As shown in Table 6, the Websites of Asian banks attracted more inlinks than their counterparts elsewhere. Since the global Yahoo! may over-represent the U.S. Websites [25], but not the Asian Websites, the finding that Asian banks sites attracted higher number of inlinks is likely to be reliable one. Note that the inlink count discrepancy between Asian banks and other banks increased from December 2008 to June 2009. These are important findings because they reflect the financial power of Asian banks, in particular the Chinese banks, as demonstrated during the recent world financial crisis. The Mergent report of the Banking industry in the Asia Pacific region [24] underlines the strong performance of Chinese banks. Some factors contributing to this include banking reforms, increases in foreign investment, stronger supervision and more intense competition. At this moment, China has emerged as a prominent player on the international stage and its banks remain stronger and healthier in comparison with many banks in the U.S. and Europe. This strength compensates for the weakness of the Japanese banking industry, which is also included in the Asian group for the analysis.

Table 6. Inlink count comparison between Asian Banks and other banks

	Number of banks	Dec 2008 median inlink count	June 2009 median inlink count
Asian Banks	9	193,000	273,000
Other banks	41	40,600	35,900

3.3 Co-link analysis

MDS maps (see Figures 1 and 2) based on the number of co-links between banks show the relative competitive positions of the banks in the study. The MDS map based on December 2008 data is shown in Figure 1. It 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 U.K.). Chinese and Japanese banks are clustered together on 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 a local level. They receive many inlinks but relatively few of them are co-links with other banks. Japanese banks are in an intermediate position with MUFG and SMBC close to U.S. banks, Mizuho to Chinese banks and Nochu and MizuhoCBK to European entities. Banks from English speaking countries (U.S., U.K., Canada and Australia) are clustered together, although national groups are also clearly identified within the cluster. Continental 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 other banks that are not clustered with their national banks (Fortis and Royal Bank of Scotland) are also in a central position. The cluster of French banks is clearly identifiable, as is the German group, although it is a bit more dispersed.

Figure 1. MDS map for based on December 2008 data
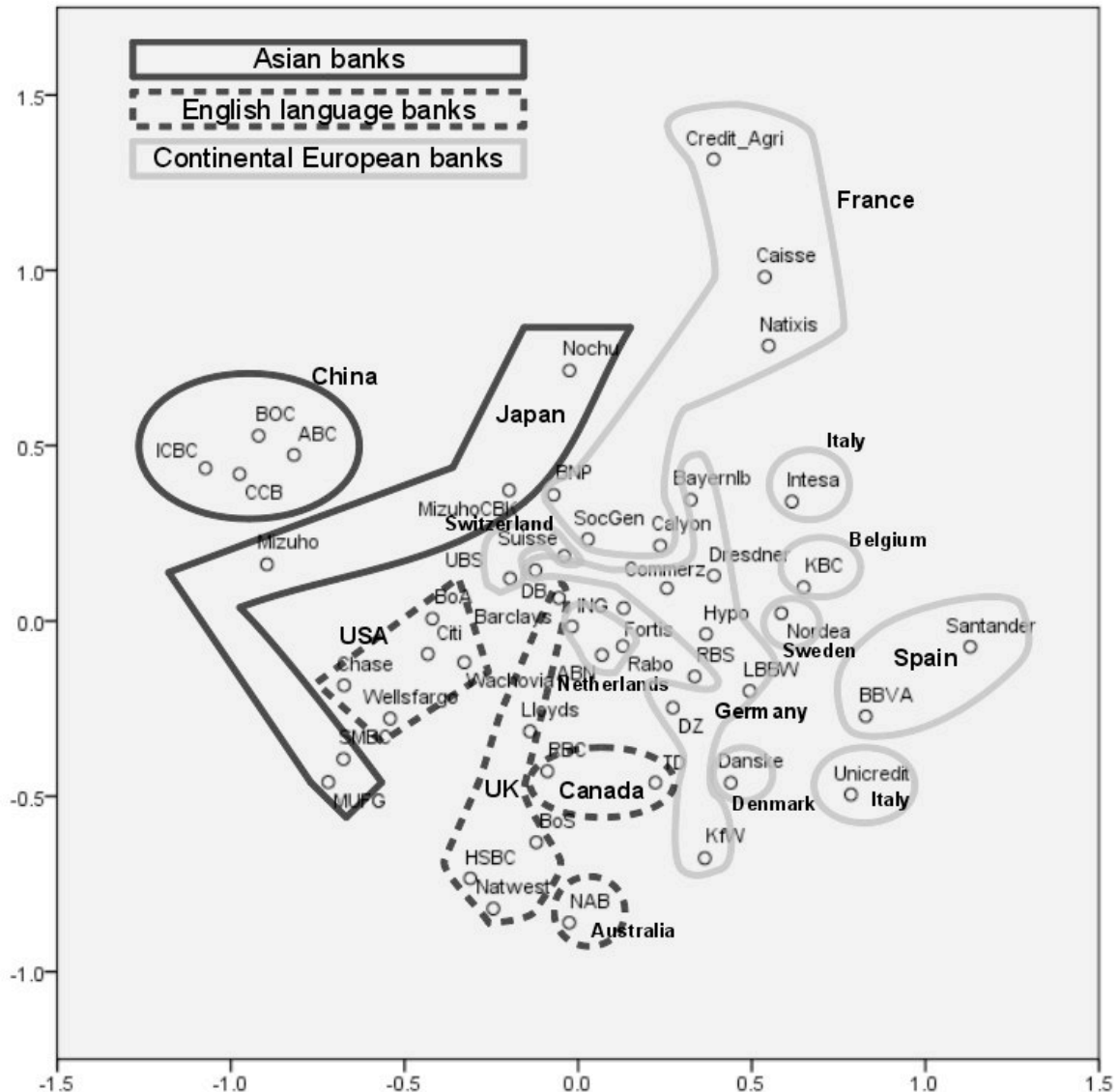
The MDS map based on June 2009 data is shown in Figure 2. To make a clear comparison with Fig. 1, banks in Fig. 2 are also clustered and labelled by country. Despite the fact that the country factor remains salient, there are some significant changes from Fig. 1 to Fig. 2 that are likely related to the evolution in the economic and financial crisis over the six month period. It is worth noting that the Chinese and Japanese banks moved apart and the U.S. banks moved between them. The Chinese banks also moved closer to the British banks compared with their position in Fig. 1. Although further monitoring and analysis is needed to find out for sure whether this reflects the changing perception the world has on the Chinese banks as the result of the world financial crisis, it is reasonable to speculate that this reflects the fact that major Asia-Pacific banks, except the Japanese banks, remain strong and well-positioned in the international context [24]. This is particularly true of the Chinese banks that are among the top in the world. Swiss banks and major European banks (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 into two groups.

Figure 2. MDS map based on June 2009 data
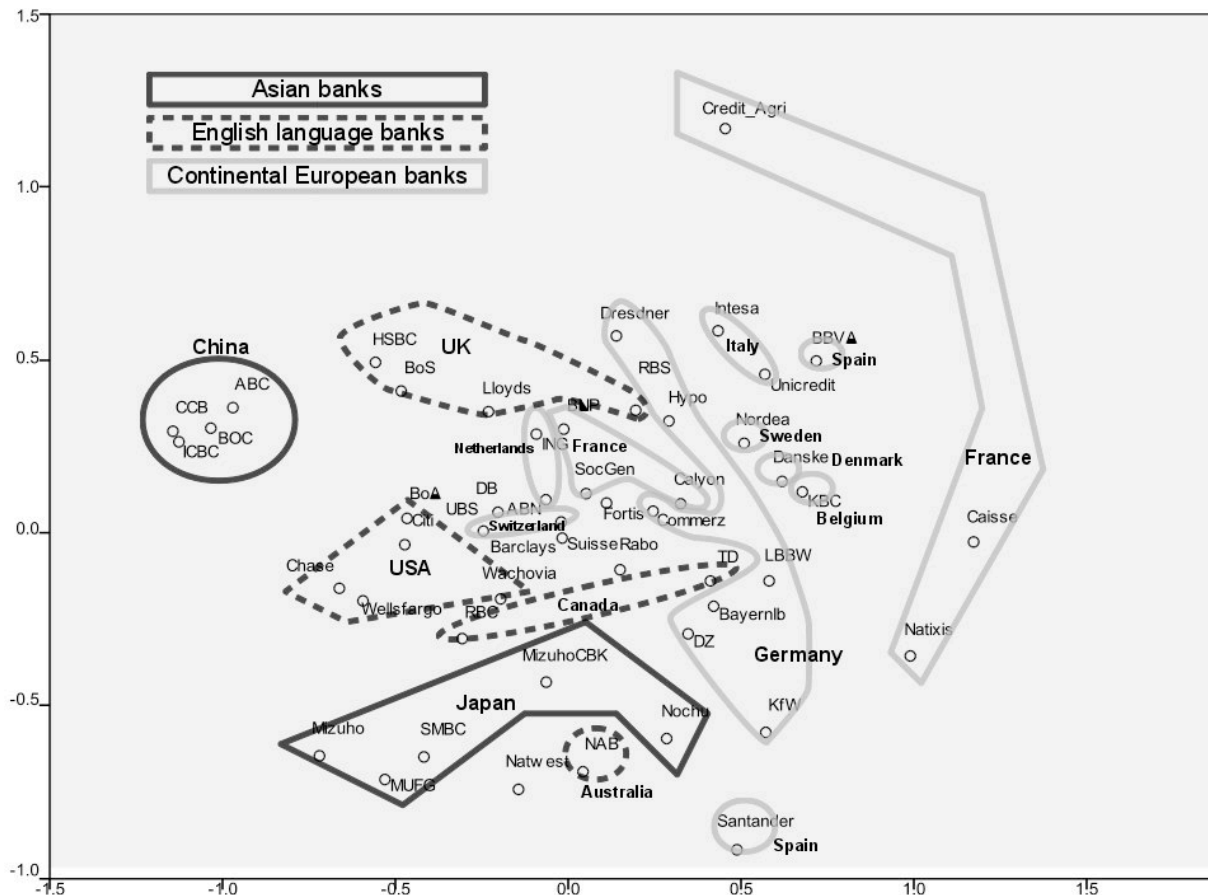
4. Conclusions and future research

The study found significant correlations between inlink counts and various financial variables, some of which have not been tested in previous studies. Findings are consistent with previous research in the same industry for a particular country [9]. This confirms that Web hyperlink data can provide useful business information even in a heterogeneous environment in which companies from multiple countries are included. A comparison of the inlink count between Asian banks and other banks showed that Asian banks received more inlinks. This is consistent with the fact that Asian banks weathered the recent world financial crisis relatively better than non-Asian banks. This finding further shows the usefulness of Web hyperlink data in gathering business information. MDS analysis based on co-links indicates that, despite the existence of main global players, the global banking industry is partially organised by regional markets. Some banks seem to be more isolated, such as the Chinese banks, for the first time period of the study. This is one potential reason why the Chinese banks didn't suffer from the recent world financial crisis to the degree that other banks did. However, data collected from the second study period revealed changing position of the Chinese banks, as they moved closer to major players such as the U.S. and U.K. banks. This move likely reflects the global financial industry's changing perception of the Chinese banks as a result of their relative resilience in the world financial crisis. Swiss banks are located at the centre of the map together with major banks, such as Barclays or Deutsche Bank, from different European countries. This echoes the central role these banks play in the world banking industry. All these findings show that Web hyperlink data

contain rich business and social information and Webometrics methods are effective in studying various Web phenomena.

There could be some limitations to the inlink and co-link analysis as a result of the inherent limitations of the search engine used for data collection. Yahoo! maintains a different database in some countries, such as China. Although this did not prove to be a problem for this particular study, as the inlink counts for Asian countries were actually higher, it could be a potential problem in other studies. The number of banks from each country is not homogeneous and this could affect the results as the main variable used in the analysis is the bank's country of origin. On the other hand, if we select banks with similar size or scale, the number of banks from each country will always be uneven because different countries have different degrees of economic power, so a choice must be made that takes various factors into consideration. It is also worth noting that the banks' Web policies are a significant issue.

Future research will focus on monitoring the changes of the MDS maps by collecting and analysing data over time. We could also expand the scale of the study to include more banks from more countries in order to find out if the findings from this study can be generalized. Although Webometrics is mainly quantitative in nature, it could be complemented by qualitative analysis such as a content analysis of the linking pages that would allow us to further our understanding of the Web linking phenomenon. Finally, it will be useful to consolidate findings from related studies to develop a systematic methodology of Webometrics for competitive intelligence.

5. References

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

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

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

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

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

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

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

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

[9] E. Romero Frías, L. Vaughan and L. Rodríguez Ariza, 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 (Unpublished proceedings XV Congreso de la Asociación Española de Contabilidad y Administración de Empresas, Valladolid, September 23-25, 2009).

[10] H. Small, 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 (1973) 265–269.

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

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

Vaughan, L. & Romero-Frías, E. (2010). "Web hyperlink patterns and the financial variables of the global banking industry". *Journal of Information Science*, 36(4): 530–541. Postprint for research purposes.

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

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

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

[16] L. Vaughan, J. Tang and J. Du, Examining the robustness of Web co-link analysis, *Online Information Review*, 33(5) (2009) 956-972.

[17] E. Romero-Frías and L. Vaughan, Financial distress of U.S. banking industry viewed through Web data. In *Proceedings of ISSI 2009 – the 12th International Conference of the International Society for Scientometrics and Informetrics* (Rio de Janeiro, Brazil, July 14-17, 2009) 206-210.

[18] Google, *Links to your site (2009)*. Available at: http://www.google.com/support/webmasters/bin/answer.py?hl=en&answer=55281 (accessed 3 June 2009).

[19] Google, *Google SOAP Search API Reference (2006)*. Available at: http://www.google.com/apis/reference.html#2_2 (accessed 3 June 2009).

[20] Live Search, *We are flattered, but... (2007)*. Available at: http://www.bing.com/community/blogs/search/archive/2007/03/28/we-are-flattered-but.aspx (accessed 3 June 2009).

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

[22] Mergent, Europe - Banking Sectors (September 2009), available at http://webreports.mergent.com (accessed 24 December 2009).

[23] J. Shaw and G. Parussini, *China bank regulator: Global regulatory cooperation very fragile (2009)*. Available at: http://www.nasdaq.com/aspx/stock-market-news-story.aspx?storyid=200910051100dowjonesdjonline000261&title=china-bank-regulatorglobal-regulatory-cooperation-very-fragile (accessed 15 October 2009).

[24] Mergent, Asia-Pacific - Banking Sectors (June 2009), available at http://webreports.mergent.com (accessed 24 December 2009).

[25] L. Vaughan and Y. Zhang, Equal representation by search engines? A comparison of Websites across countries and domains, *Journal of Computer-Mediated Communication*, 12(3) (2007), available at http://jcmc.indiana.edu:80/vol12/issue3/vaughan.html

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	U.K.	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	U.K.	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	U.S.A.	223,000	144,000
BoA	Bank of America NA	http://www.bankofamerica.com	U.S.A.	339,000	304,000
Citi	Citibank NA	http://www.citibank.com	U.S.A.	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	U.K.	13,700	9,390
HSBC	HSBC Bank plc	http://www.hsbc.co.uk	U.K.	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	U.K.	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	U.S.A.	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	U.K.	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	U.S.A.	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