Principal component analysis of financial statements. A compositional approach

CARRERAS SIMÓ, MIQUEL
Department of Business Studies
University of Girona
Correo electrónico: miquel.carreras@udg.edu

COENDERS, GERMÀ
Department of Economics, Faculty of Economics and Management
University of Girona
Correo electrónico: germa.coenders@udg.edu

ABSTRACT

Financial ratios are often used in principal component analysis and related techniques for the purposes of data reduction and visualization. Besides the dependence of results on ratio choice, ratios themselves pose a number of problems when subjected to a principal component analysis, such as skewed distributions. In this work, we put forward an alternative method drawn from compositional data analysis (CoDa), a standard statistical toolbox for use when data convey information about relative magnitudes, as financial ratios do. The method, referred to as the CoDa biplot, does not rely on any particular choice of financial ratio but allows researchers to visually order firms along the pairwise financial ratios for any two accounts. Non-financial magnitudes and time evolution can be added to the visualization as desired. We show an example of its application to the top chains in the Spanish grocery retail sector and show how the technique can be used to depict strategic management differences in financial structure or performance, and their evolution over time.

Análisis en componentes principales de los estados financieros. Un enfoque composicional

RESUMEN

Las ratios financieras se utilizan a menudo en el análisis en componentes principales y técnicas relacionadas con el fin de reducir y visualizar los datos. Además de la dependencia de los resultados de la elección de las ratios, las ratios en sí plantean una serie de problemas cuando se someten a un análisis de componentes principales, por ejemplo, distribuciones asimétricas. En este trabajo, presentamos un método alternativo que proviene del análisis de datos composicionales (CoDa), una caja de herramientas estadística estándar para usar cuando los datos contienen información sobre magnitudes relativas, como lo hacen las ratios financieras. El método, conocido como el biplot CoDa, no se basa en una elección particular de ratios financieras, sino que permite a los investigadores ordenar visualmente las empresas a lo largo de las ratios financieras entre cualesquiera pares de cuentas. Las magnitudes no financieras y la evolución temporal se pueden agregar a la visualización como se desee. Mostramos un ejemplo de su aplicación a las principales cadenas de supermercados españolas y mostramos cómo la técnica puede utilizarse para describir las diferencias de gestión estratégica en la estructura o el rendimiento financieros, y su evolución en el tiempo.

1. Introduction.

Financial ratios have been used for a variety of purposes, ranging from bankruptcy prediction and option pricing to management performance assessment and strategic assessment (e.g., Barnes, 1987; Blanco-Oliver et al., 2016; Caro et al., 2017; Chen & Shimerda, 1981; Horrigan, 1968; McGee & Thomas, 1986). As regards the latter purposes, since the last half of the nineteenth century, financial ratio analysis has been a common methodology for assessing firm management (Horrigan, 1968). The original DuPont model based on financial ratios developed in 1918 has become a universally used approach to analyse the relationship between management and firm performance. Since then, scholars and educators alike have acknowledged the role that financial ratios play in the formulation and implementation of realistic and ultimately effective strategies (Norman, 2018). Ross et al. (2003) considered that one the main advantages of financial ratios is their ability to evaluate the company's position compared to its main competitors. This is an essential aspect to address in the strategic management of a company. Strategic management is a process that establishes the courses of action a company will follow to achieve its objectives (Grant, 2008; Smith, 2005). Techniques such as strengths, weaknesses, opportunities and threats (SWOT) analysis, the balanced scorecard, and conceptual approaches such as competitive advantage, resource-based analysis or strategic groups, are widely acknowledged and used in strategic management and business analysis. Their application requires having quantitative information that makes it possible to assess the situation of a company and its competitors, as well as project future strategic scenarios. Financial ratios applied to general-purpose financial statements and related data are the core of useful estimates and inferences in business analysis. In fact, financial ratios make it possible to assess business strategy and relate it to performance measures (Allen & Helms, 2006; Banker et al., 2014; Hoque, 2004). Despite the widespread and frequent use of financial ratios as a business analysis tool, it has been hampered by several long-known shortcomings related to financial-ratio asymmetry and redundancy. Most ratios are distributed between zero and infinity and thus make fully symmetric distributions impossible to achieve (Deakin, 1976). Besides its statistical implications, asymmetry arises from an uneven treatment of differences in the numerator and the denominator of the ratio (Frecka & Hopwood, 1983), meaning that permuting them changes the results of statistical analyses (Linares-Mustarós et al., 2018).

Redundancy necessarily arises from the fact that there are many more financial ratios in common use than there are accounts from which these ratios are computed. Often, redundancy occurs to such an extent that *"there is no absolute test for the importance of variables"* (Barnes, 1987, p. 455) and *"to identify those ratios which contain complete information about a firm while minimising duplication cannot be achieved purely by logic"* (Barnes, 1987, p. 456). In extreme cases, there is an exact dependency between ratios. Chen & Shimerda (1981) put an example of four exactly redundant ratios: net worth to total debt, total debt to net worth, net worth to total assets and total debt to total assets. Just one of them carries the same information as the whole set of four.

Financial ratio analysis constitutes a case of genuine interest to researchers and professionals with regard to relative rather than absolute magnitudes of accounts in financial statements. In other scientific fields, there is a well-developed toolbox for analysing the relative importance of magnitudes, known as Compositional Data Analysis (CoDa- Aitchison's, 1982, 1986; van den Boogaart & Tolosana-Delgado, 2013; Pawlowsky-Glahn et al., 2015). Among other features, CoDa treats magnitudes (i.e., account values) in a symmetric fashion in such a way that results depend only on the selected accounts of interest but not on any particular set of ratios, let alone on numerator and denominator permutation. CoDa also tends to reduce redundancy and hence increase parsimony by acknowledging the fact that no analysis will require more variables than there are account magnitudes to be compared.

CoDa has already been successfully applied with the purpose of clustering firms with similar financial statement structures (Linares-Mustarós et al., 2018). To the best of our knowledge, that is the only CoDa application to financial statement analysis to date. The purpose of the present article is to introduce a data visualization and data reduction tool for financial statement analysis based on CoDa. With this tool firms can be mapped in a two-dimensional space, which allows ready appraisal of their

ordering according to the ratios between any two accounts, thus providing a powerful visual aid for strategic assessment.

In this article, we first provide an overview of the basics in CoDa. Then, we present a common data visualization and reduction tool called CoDa biplot, which can be used to map the main accounts in order to visually appraise the strategic or managerial differences among firms. Some interpretational cues specific to financial ratio analysis are highlighted. Finally, we provide an illustration characterizing the business model evolution of the largest chains in the Spanish grocery retailing sector, from the onset of the financial crisis in 2008 up to 2015.

2. Compositional data analysis.

2.1. Definition and purpose.

Compositional Data analysis (CoDa) is the standard statistical method used when data carry only information about the relative importance of non-negative parts of a whole. The CoDa tradition started with Aitchison's seminal work (1982, 1986) on chemical and geological compositions, in which only the proportion of each part or component is of interest, since absolute amounts are irrelevant and only telling with regard to the size of the chemical or soil sample (e.g., Buccianti et al., 2006). Nowadays, CoDa spans almost all of the hard sciences and has started to be used in several management fields. Besides Linares et al. (2018), examples include crowdfunding (Davis et al., 2017), financial markets (Ortells et al., 2016; Wang et al., 2019), investment portfolios (Belles-Sampera et al., 2016; Boonen et al., 2019; Glassman & Riddick, 1996), municipality budgets (Voltes-Dorta et al., 2014), product portfolios (Joueid & Coenders, 2018), market segmentation (Ferrer-Rosell & Coenders, 2018; Ferrer-Rosell et al., 2016a), market share (Morais et al., 2018), advertisement (Mariné-Roig & Ferrer-Rosell, 2018), consumer research (Ferrer-Rosell et al., 2015; Ferrer-Rosell et al., 2016b; Vives-Mestres et al., 2016a), quality management (Vives-Mestres et al., 2014; 2016b), organizational culture (Van Eijnatten, et al., 2015), and management education (Batista-Foguet et al., 2015; Mateu-Figueras et al., 2016).

In the last three decades, CoDa has provided a standardized toolbox for statistical analyses whose research questions concern the relative importance of magnitudes. Dedicated user-friendly software is available to this end (Van den Boogaart & Tolosana-Delgado, 2013; Greenacre, 2018; Palarea-Albaladejo & Martín-Fernández, 2015; Templ et al., 2011; Thió-Henestrosa & Martín-Fernández, 2005), as well as accessible handbooks (Van den Boogaart & Tolosana-Delgado, 2013; Filzmoser et al., 2018; Greenacre, 2018; Pawlowsky-Glahn & Buccianti, 2011; Pawlowsky-Glahn et al., 2015).

Compositional analysis (Barceló-Vidal & Martín-Fernández, 2016) has recently been coined as a term to emphasize the fact that what is ultimately compositional are not the data but the analysis and research objectives centred on the relative importance of non-negative magnitudes. Interesting applications of CoDa to non-negative data that do not represent parts of any whole can be found in Ortells et al. (2016) and Azevedo-Rodrigues et al., (2011). This is the case with financial statement analysis, in which, for instance, sales and assets are not parts of any whole and the asset turnover ratio compares the magnitudes of both in relative terms. Even non-negative non-financial magnitudes can be and indeed are included, such as number of employees.

Compositional data are represented by a positive vector in a D–dimensional real space, which conveys information about the relative size of its components:

$$\mathbf{x} = (x_1, x_2, ..., x_D) \in R_+^D \ , \ \text{with} \ \ x_j > 0 \ \ \text{for all} \ j = 1, 2, ..., D, \qquad (1)$$

where D is the number of components, also referred to as parts. In our case, parts are non-negative accounts in financial statements or even other non-negative management magnitudes. This means, for instance, that one should use the non-negative constituents of working capital (current assets and current

liabilities) rather than working capital itself, that one should use revenues and costs rather than profit, that one should use assets and liabilities rather than net worth, and so on.

2.2. Transformations, variance and association.

The most common CoDa approach is to express an original compositional vector of D parts into logarithms of ratios (Aitchison, 1986; Egozcue et al., 2003). *Log ratios* are unbounded and tend to meet the distributional assumptions of classical statistical models, such as normality (Aitchison, 1982; Pawlowsky-Glahn et al., 2015). This notwithstanding, for the purposes of this article the main arguments for log ratios are that they constitute a natural way of distilling information about relative size of parts, they form the basis for defining compositional association and variance in a meaningful way, and, involving ratios, they are coherent with financial statement analysis practice. Logs of financial ratios had already been suggested in the financial ratio literature as a means of reducing asymmetry (e.g. Cowen & Hoffer, 1982; Deakin, 1976; Sudarsanam & Taffler, 1995).

Log ratios may be computed among all possible pairs of parts:

$$\ln \frac{x_j}{x_k} \quad \text{with } j,k = 1,2,...,D; \, j \neq k. \tag{2}$$

Positive values mean that x_j is larger than x_k. Negative values show the opposite. A zero log ratio implies equality of both magnitudes, exactly in the same way as a unit standard ratio. If x_j contains quick assets and x_k contains current liabilities, interpretation parallels that of the quick ratio. A log ratio is symmetric in the sense that its range is from minus infinity to plus infinity. It is also symmetric in the sense that permuting the numerator and denominator only affects the log ratio sign, but not its value. Furthermore, if one of the parts being compared is close to zero, it may lead to an outlying standard financial ratio when placed in the denominator and to a typical ratio when placed in the numerator. For log ratios placement makes no difference.

Log ratios may also be computed between each part and the geometric mean the remaining D–1 parts, in the so-called *centred log ratios*. They thus indicate the relative dominance of an account in the overall financial statement structure. $(D–1)/D$ is a scaling constant:

$$\frac{D-1}{D} \ln \left(\frac{x_j}{\sqrt[D-1]{x_1 x_2 ... x_{j-1} x_{j+1} ... x_D}} \right) \text{with } j = 1,2,...,D. \tag{3}$$

There are alternative interpretations and expressions of centred log-ratios (Filzmoser et al., 2018; Pawlowsky-Glahn et al., 2015). Other ways of computing log ratios in financial analysis are described in Linares-Mustarós et al. (2018).

Total variance in a compositional data set is expressed by the sum of variances of all centred log ratios:

$$\sum_{j=1}^{D} Var \left(\frac{D-1}{D} \ln \left(\frac{x_j}{\sqrt[D-1]{x_1 x_2 ... x_{j-1} x_{j+1} ... x_D}} \right) \right). \tag{4}$$

Association is understood as proportionality between pairs of accounts is (Lovell et al., 2015). The same pairwise log ratios (2) and their variances are computed:

$$Var \left(\ln \frac{x}{x_k} \right) \quad \text{with } j,k = 1,2,...,D; \, j \neq k. \tag{5}$$

These variances can be arranged in a symmetric matrix with components (i.e., accounts) defining both D rows and D columns, with the same layout as a correlation matrix. This is the so-called variation matrix and has zero elements in the diagonal. $\text{Var}(\ln(x_j/x_k))$ is zero when x_j and x_k behave perfectly proportionally (e.g., firms having one account of double size also have the other of double size), which corresponds to perfect positive association (Lovell et al., 2015). The further $\text{Var}(\ln(x_j/x_k))$ is from zero, the lower the association. There is no clearly defined threshold representing no association, just as there is no upper bound representing perfect negative association, so that values in the matrix are assessed comparatively. It can be shown that the sum of elements in the variation matrix is $2D$ times the total variance (4).

2.3. Zero and other irregular components.

As with other types of data, compositions require some data preprocessing to deal with problems such as missing information and outliers.

To begin with, the accounts of interest may contain no zero values in order for log ratios to be computed (e.g. Martín-Fernández et al., 2011). The same holds for standard financial ratio analysis regarding the account in the denominator. Unlike the case in standard financial ratio analysis, CoDa includes an advanced toolbox for zero imputation prior to log ratio computation, with Martín-Fernández et al. (2011) and Palarea-Albaladejo & Martín-Fernández (2008) being key references. This makes financial statement analysis possible even when some accounts of interest equal zero. An advanced formal imputation method is the modified EM algorithm in Palarea-Albaladejo & Martín-Fernández (2008).

CoDa also has implications for outlier detection. Given that components cannot be considered in isolation, multivariate outlier detection methods are called for. Once compositions have been transformed into log ratios, squared Mahalanobis distances between each composition and the overall mean can be computed. Under multivariate normality, these squared distances follow a χ^2 distribution with $D-1$ degrees of freedom. An appropriate percentile for this distribution (e.g., the 99.9 percentile) can be used as cut-off criterion for outlier detection.

3. Compositional principal component analysis and the CoDa biplot.

Based on the seminal work by Pinches et al. (1973), the application of factor analysis and related data reduction methods has spread as a means of summarizing the extensive information offered by the many common financial ratios and determining patterns of financial structure and performance. These methods include, among others, principal component analysis (e.g., Chen & Shimerda, 1981; Cowen & Hoffer, 1982; Dimitropoulos et al., 2010; Martín-Oliver et al., 2017; Sanz et al., 2018; Sharma et al., 2016; Yoshino & Taghizadeh-Hesary, 2015; Yoshino et al., 2016) and factor analysis (e.g., Lukason & Laitinen, 2016; Pinches et al., 1973; Yap et al., 2014).

CoDa also has its own multivariate statistical methods, which are in many cases closely related to standard methods used for log ratio transformations. In this article, we deal with compositional principal component analysis. Aitchison (1983) was the first to extend principal component analysis to the compositional case. The extension boils down to submitting centred log ratios (3) to an otherwise standard principal component analysis based on a covariance matrix, and adapting interpretation to take into account the fact that the information carried by the data refers to relative rather than absolute importance of magnitudes. Together with Gabriel's (1971) *biplot*, which jointly represents cases and variables in a principal component analysis, Aitchison's developments served as a basis for Aitchison & Greenacre (2002) introducing CoDa biplots.

A CoDa biplot can be understood as the most accurate data visualization of a compositional dataset in two dimensions. As in standard principal component analysis, overall biplot accuracy can be assessed from the percentage of variance (4) explained by the two first principal components. There are

several types of biplots. The most interesting type for financial statement analysis purposes is the so-called *covariance biplot*, which optimizes the representation of the variation matrix (5) among the selected accounts and non-financial magnitudes. Accounts and other magnitudes appear as rays emanating from a common origin, which represents a firm with log ratios equal to the sample average. Firms appear as points. The interpretation of the covariance biplot is explained below (see Aitchison & Greenacre, 2002; Van den Boogaart & Tolosana-Delgado, 2013; Pawlowsky-Glahn et al., 2015 for further details). The two key interpretational elements are the rays representing accounts and other magnitudes, and the links between the vertices of a pair of rays:

- Lengths of the links between the vertices of the rays of two accounts or magnitudes are approximately proportional to the square root of the variance of their corresponding pairwise logratio (5). Accounts which behave proportionally for all firms appear close together. It must be noted that, unlike in standard principal component analysis, distances between vertices are used rather than angles between rays.

- The orthogonal projection of all firms along the direction defined by a ray shows an approximate ordering of the relative size of that account or magnitude for all firms. This coincides with the interpretation in standard principal component analysis, the only difference being that relative rather than absolute size is considered, as implied in (3).

- The orthogonal projection of all firms along the direction defined by the link between the vertices of a pair of rays shows an approximate ordering of firms according to the standard financial ratio between the corresponding two accounts. In this way, the CoDa biplot is also a visual representation of any of the $D(D-1)/2$ possible financial ratios computed from any two accounts or magnitudes, although only long links showing high variance pairwise log ratios tend to lead to informative directions. The ability to visually interpret ratios between any two magnitudes is of great interest in financial statement analysis. This interpretation is specific to CoDa biplots, and differs from that of standard principal component analysis. Unlike principal component analysis of standard financial ratios, the choice of ratios of interest does not need to be made in advance and hence does not influence the analysis outcome.

- The cosine of the angles between two links corresponding to the financial ratios between two pairs of accounts shows the approximate correlation between the two corresponding log ratios. For example, parallel links show log ratios with approximate correlations equal to 1 or –1. Orthogonal links show log ratios with approximate 0 correlation.

The use of compositional principal components and the CoDa biplot is not limited to visualizing the composition but is also appropriate for data reduction and summarization purposes. The first few principal component coordinates explaining most of the variance can be used as variables in further statistical analyses, exactly in the same way as in standard principal component analysis.

4. A case study: the largest retail chains in the Spanish grocery sector (2008-2015).

This example focuses on the largest retail chains in the Spanish grocery sector. Data are taken from the largest chains in 2015 (a total of 28 chains with a net income of over €200 million) using the SABI (Iberian Balance sheet Analysis System) database, developed by INFORMA D&B in collaboration with Bureau Van Dijk. These data correspond to annual accounts presented in official records and other non-financial information from 2008 to 2015. Four chains were ruled out due to missing or faulty information for the total period considered. The final sample size was therefore 24 retail chains, i.e., 192 observations over 8 years: Ahorramás, AlCampo, Alimerka, BonPreu, Caprabo, Carrefour, Cecosa Supermercados, Condis, Consum, Dia, Distribuciones Froiz, Eroski, HiperCor, Hiper Usera, JuanFornes, Lidl, Mercadona, Semark, Supermercados Champion, Supercor, Supeco Maxor, Superercados Sabeco, Unión Detallistas Españoles, Vego Supermercados. The database was accessed on 28/7/2017.

This sector is especially interesting for analysis due to the pressures and changes it has had to face during the last economic crisis. On the one hand, we must consider a very adverse international context at the beginning of the crisis with an inflationary process in the food and energy markets (OECD, 2013) and on the other, the factors characteristic of the Spanish economy: demographic stagnation, a fall in the purchasing power of domestic economies, decrease in household size, changes in purchasing habits (Pablos et al., 2013) and an increase in VAT. Faced with this complex scenario, large grocery distribution chains have been forced to react and modify or intensify their strategic approaches, and in so doing they have contributed to shaping the sector's recent evolution (OECD, 2013). The observed trends for this sector in Spain are not isolated, coinciding with those observed for the vast majority of EU countries (European Union, 2016), although they have presented greater intensity in the Spanish case.

The non-negative financial and non-financial magnitudes used in this case study include:

x_1=Fixed assets (FA)
x_2=Inventory (I)
x_3=Quick assets (QA)
x_4=Long term liability (LTL)
x_5=Short term debt (STD)
x_6=Accounts payable (AP)
x_7=Number of employees (E)
x_8=Net Sales (NS)
x_9=Costs of goods sold (CGS)
x_{10}=Labour costs (CL)
x_{11}=Costs of external services (CES)
x_{12}=Asset depreciation and amortization (AAD)
x_{13}=Other non-finance costs (CONF)
x_{14}=Finance costs (CF)

These magnitudes have been selected because they constitute the basis for a wide array of common financial and management ratios frequently used in the retail sector (Evans & Mathur, 2014), which belong to three main categories: solvency and liquidity, operating efficiency and profitability. Table 1 shows examples of relevant financial ratios that might be computed from x_1 to x_{14}. The use of CoDa does not favour one ratio over any other, as only the data transformed as centred log ratios (3) are used in the analysis. This ratio list is not closed; it could incorporate any possible ratio computed from x_1 to x_{14}. Thus, it is provided merely as an example in order to show which type of information is carried by x_1 to x_{14}.

Table 1. A sample financial ratio list, which might be computed from x_1 to x_{14}.

Ratios	Formal expression
Solvency and liquidity	
Current ratio	(I+QA)/(STD+AP)
Quick ratio	QA/(STD+AP)
Current liabilities to net worth	(AP+STD)/(FA+I+QA–LTL–STD–AP)
Current liabilities to inventory	(AP+STD)/I
Total liabilities to net worth	(AP+STD+LTL)/(FA+QA+I–LTL–STD–AP)
Fixed assets to net worth	FA/(FA+QA+I–LTL–STD–AP)
Temporary structure of the liabilities	(STD+AP)/(STD+AP+LTL)
Average financial cost of liabilities	CF/(STD+AP+LTL)
Operating efficiency	
Net sales to inventory	NS/I
Net sales to assets (Asset turnover)	NS/(FA+QA+I)
Net sales to fixed assets (Fixed asset turnover)	NS/FA
Asset depreciation and amortization over fixed assets	AAD/FA
Net sales to net working capital	NS/[(QA+I)–(AP+STD)]
Accounts payable to net sales	AP/NS
Accounts payable turnover	(AP/CGS)
Fixed assets per employee	FA/E
Net sales per employee (Productivity)	NS/E
Labour cost to operating costs	CL/(CGS+CL+CES+AAD)
Gross margin	(NS–CGS)/NS
Profitability	
Return on sales (EBIT margin)	[NS–(CGS+CL+CES+AAD+CONF)]/NS
Return on assets	[NS–(CGS+CL+CES+AAD+CONF)]/(FA+I+QA)
Return on equity	[NS–(CGS+CL+CES+AAD+CONF+CF)]/ [(FA+I+QA)–(LTL+STD+AP)]

Source: Own elaboration.

All analyses were performed with the R packages compositions (Van den Boogaart, & Tolosana-Delgado, 2013) and zCompositions (Palarea-Albaladejo & Martín-Fernández, 2015). Component x_{14}=CF had two zero values, which were imputed with the modified EM algorithm (Palarea-Albaladejo & Martín-Fernández, 2008) as implemented in the lrEM command with default options and the minimum non-zero observed value as detection limit. Eight outliers with squared Mahalanobis distance to the centre above the 99.9 percentile of the χ^2 distribution were removed. The final sample size was thus *n*=184. The acomp, clr and princomp commands were used in the principal component analysis. However, once centred log ratios (3) are computed, any standard software handling covariance-based principal component analysis could be used with identical results. Subsequently, the biplot command was used, whose default options return the covariance biplot.

5. Results.

The variation matrix in Table 2 shows some cells with very low values, in other words, ones displaying strong proportionality. Proportionality between number of employees (E) and labour costs (CL) suggests average wages to be very similar across chains and over time. Proportionality among net sales (NS), cost of goods sold (CGS), labour costs (CL), and costs of external services (CES) suggests that chains have very similar operating margins and cost structure.

Table 2. Variation matrix, and centred log ratio variances.

|	FA	I	QA	LTL	STD	AP	E	NS	CGS	CL	CES	AAD	CONF	CF
						Variation matrix								
FA	0.00	0.43	1.19	1.49	0.35	1.66	0.41	0.42	0.48	0.37	0.25	0.32	0.49	2.44
I	0.43	0.00	0.80	2.21	0.40	0.98	0.25	0.10	0.11	0.15	0.13	0.26	0.61	2.92
QA	1.19	0.80	0.00	2.98	1.11	1.06	0.99	0.52	0.50	0.72	0.79	1.00	1.42	3.60
LTL	1.49	2.21	2.98	0.00	1.88	3.01	2.54	2.38	2.46	2.41	2.03	2.33	2.16	1.54
STD	0.35	0.40	1.11	1.88	0.00	1.55	0.52	0.44	0.46	0.49	0.36	0.52	0.76	2.33
AP	1.66	0.98	1.06	3.01	1.55	0.00	1.45	0.87	0.84	1.05	1.19	1.46	1.68	3.06
E	0.41	0.25	0.99	2.54	0.52	1.45	0.00	0.26	0.32	0.08	0.18	0.22	0.81	3.31
NS	0.42	0.10	0.52	2.38	0.44	0.87	0.26	0.00	0.01	0.10	0.10	0.22	0.67	3.01
CGS	0.48	0.11	0.50	2.46	0.46	0.84	0.32	0.01	0.00	0.14	0.14	0.27	0.72	3.04
CL	0.37	0.15	0.72	2.41	0.49	1.05	0.08	0.10	0.14	0.00	0.10	0.16	0.67	3.14
CES	0.25	0.13	0.79	2.03	0.36	1.19	0.18	0.10	0.14	0.10	0.00	0.16	0.59	2.84
AAD	0.32	0.26	1.00	2.33	0.52	1.46	0.22	0.22	0.27	0.16	0.16	0.00	0.65	3.20
CONF	0.49	0.61	1.42	2.16	0.76	1.68	0.81	0.67	0.72	0.67	0.59	0.65	0.00	3.21
CF	2.44	2.92	3.60	1.54	2.33	3.06	3.31	3.01	3.04	3.14	2.84	3.20	3.21	0.00
						Centred log ratio variances								
	0.20	0.14	0.66	1.57	0.27	0.89	0.28	0.12	0.15	0.16	0.10	0.24	0.50	2.16

Source: Own elaboration.

Table 3 and Figure 1 show two dimensions to be appropriate for a compositional principal component analysis and thus the two-dimensional biplot to accurately represent centred log ratio variances (4) and the variation matrix (5).

Table 3. Percentages of explained variance by components.

| Comp.	1	2	3	4	5	6	7	8	9	10	11	12	13
%	49.6	19.5	9.4	7.2	5.5	3.1	2.0	1.5	1.1	0.5	0.4	0.2	0.0
Cum %	49.6	69.1	78.5	85.7	91.2	94.4	96.3	97.9	98.9	99.5	99.8	100.0	100.0

Source: Own elaboration.

Figure 1. Principal component screeplot.
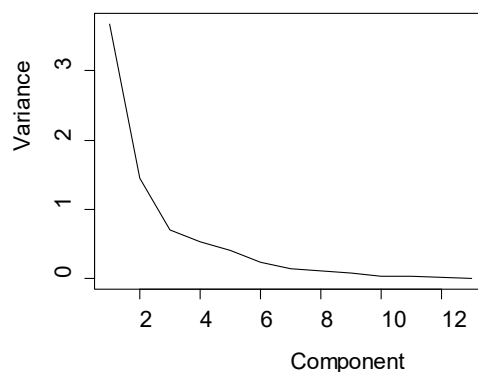
Source: Own elaboration.

Figure 2 shows the covariance biplot with only part rays (only 2005 rays are shown). As expected, magnitudes related to net sales and direct costs (NS, AAD, CES, CGS, CL) and thus to operating margins appear close together. Heterogeneity among chains mostly lies in solvency, liquidity and turnover. The most leveraged chains will appear to the right of the graph (high financial costs –CF– and long term liabilities –LTL–) relative to their activity volume (sales –NS–) and non-financial operating costs –AAD, CES, CGS, CL–). The upper part of the biplot represents chains with high fixed assets (FA) and overheads (CONF) relative to the quick assets (QA) and liabilities (AP) related to core activity (NS).
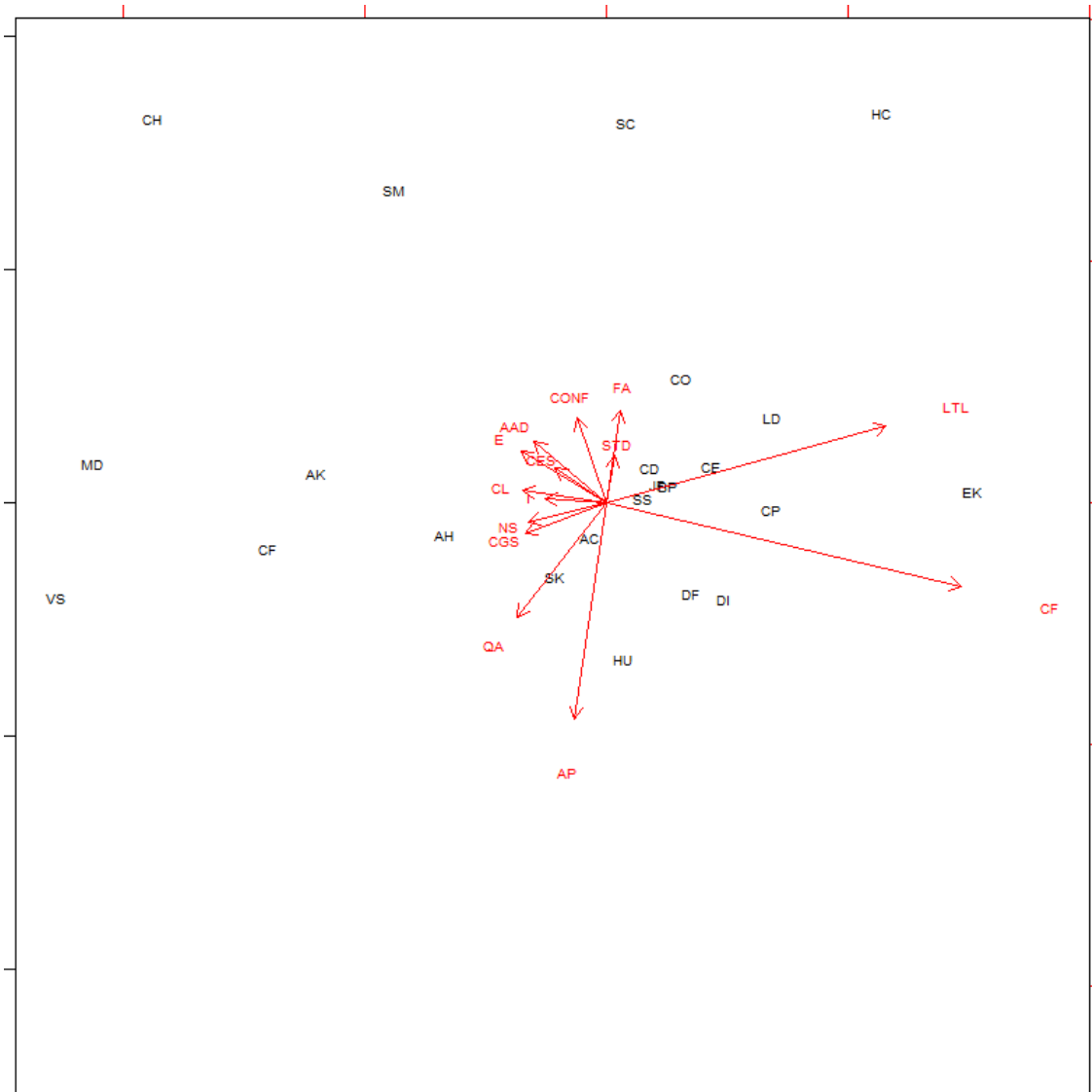
Figure 2. Covariance biplot.
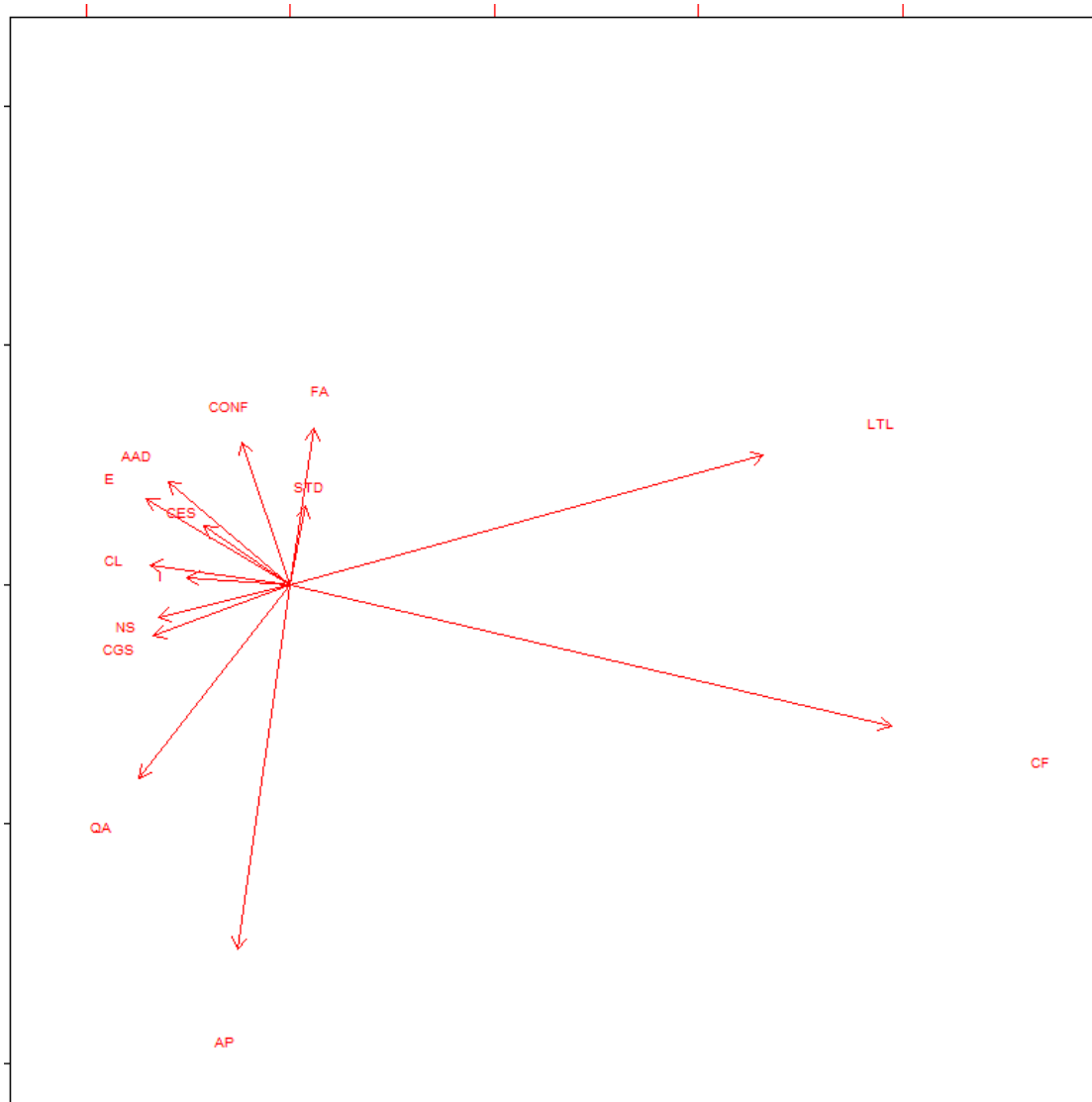
Source: Own elaboration.

Figure 3 shows the covariance biplot with both rays and chains (only 2015 data shown). In spite of the fact that all observations have been used for its computation, only 2015 values are shown for the sake of simplicity.
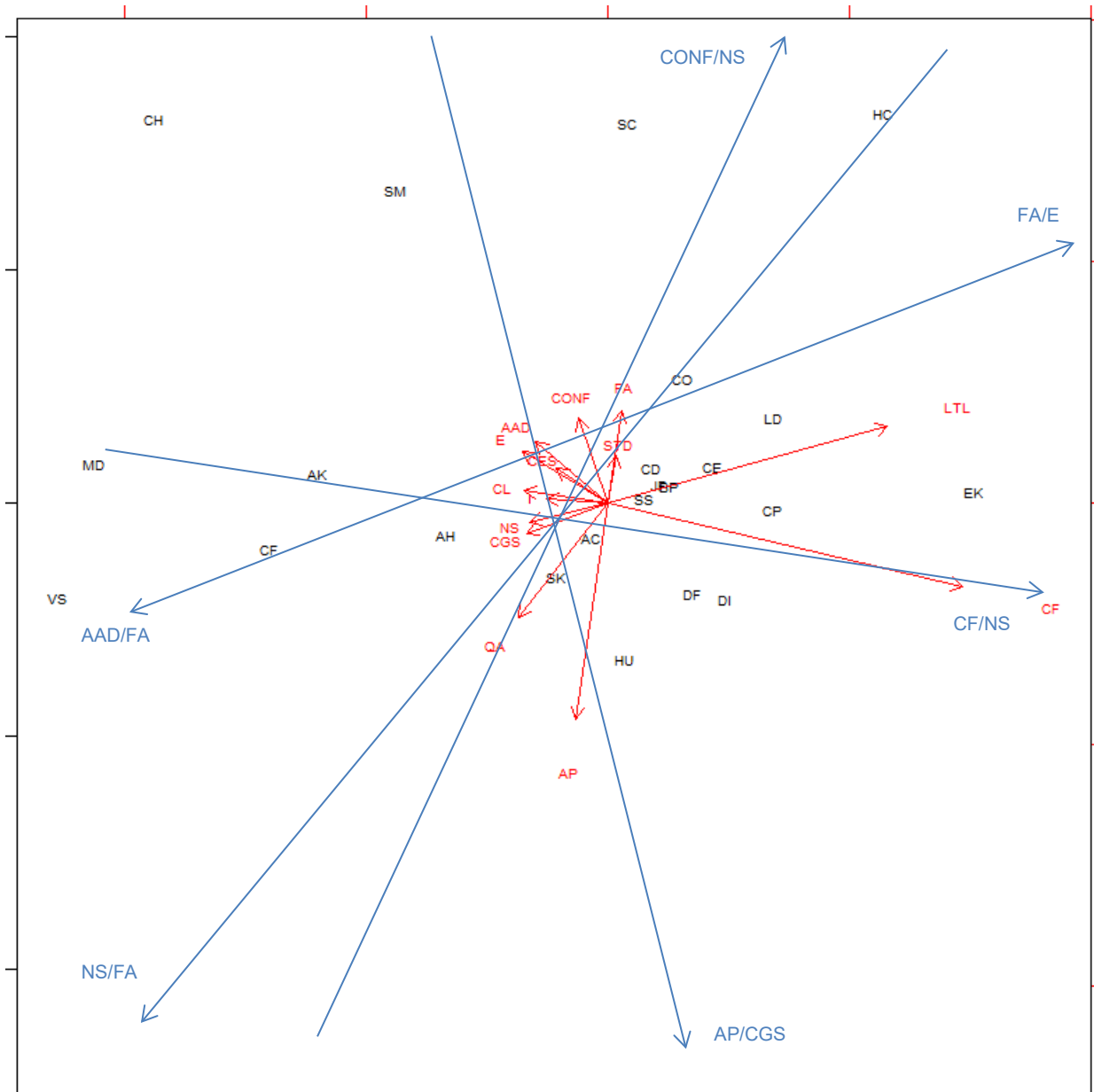
Figure 3. Full covariance biplot.

Ahorramás (AH), AlCampo (AC), Alimerka (AK), BonPreu (BP), Caprabo (CP), Carrefour (CF), Cecosa Supermercados (CE), Condis (CD), Consum (CO), Dia (DI), Distribuciones Froiz (DF), Eroski (EK), HiperCor (HC), Hiper Usera (HU), JuanFornes (JF), Lidl (LD), Mercadona (MD), Semark (SK), Supermercados Champion (CH), Supercor (SC), Supeco Maxor (SM), Superercados Sabeco (SS), Unión Detallistas Españoles (UD, outlier in 2015), Vego Supermercados (VS).

Source: Own elaboration.

Figure 4 shows the covariance biplot with the directions defined by some informative financial ratios between two magnitudes (only 2015 data shown). Note that we have not represented the ratios of net sales over inventory (NS/I) or net sales per employee (NS/E), which have very low variances given the proximity of their pairs of vertices. The represented ratios are:

- Fixed assets turnover (NS/FA).
- Share of overheads in the margin structure (CONF/NS).
- Share of financial costs in the margin structure (CF/NS).
- Fixed asset per employee (FA/E).
- Accounts payable turnover (AP/CGS).
- Asset depreciation and amortization over fixed assets (AAD/FA).

Figure 4. Covariance biplot with the directions defined by some key ratios.

Ahorramás (AH), AlCampo (AC), Alimerka (AK), BonPreu (BP), Caprabo (CP), Carrefour (CF), Cecosa Supermercados (CE), Condis (CD), Consum (CO), Dia (DI), Distribuciones Froiz (DF), Eroski (EK), HiperCor (HC), Hiper Usera (HU), JuanFornes (JF), Lidl (LD), Mercadona (MD), Semark (SK), Supermercados Champion (CH), Supercor (SC), Supeco Maxor (SM), Superercados Sabeco (SS), Unión Detallistas Españoles (UD, outlier in 2015), Vego Supermercados (VS).

Source: Own elaboration.

The angles between the links show the log fixed assets turnover (NS/FA) to be negatively correlated with the log share of overheads in the margin structure (CONF/NS) but to have almost no correlation with the log share of financial costs in the margin structure (CF/NS). Log fixed asset per employee (FA/E) also has a very low correlation with log accounts payable turnover (AP/CGS). Any other ratio of interest to the researcher and the correlation with any other ratio could be represented and interpreted in the same manner.

In spite of the fact that all observations have been used for its computation, only 2015 values are shown in Figure 4 for the sake of simplicity. The orthogonal projection of all chains along the direction defined by a ratio shows an approximate ordering of chains according to the ratio value. For instance, Hiper Usera (HU) has the highest accounts payable turnover (AP/CGS) and Supermercados Champion

(CH) the lowest. In the same vein, Eroski (EK), and HiperCor (HC) have the highest fixed asset per employee (FA/E), and Vego Supermercados (VS) the lowest. Vego Supermercados (VS) has the highest fixed assets turnover (NS/FA) and HiperCor (HC) the lowest. Hipercor (HC) has the largest share of overheads in the margin structure (CONF/NS), Eroski (EK) the largest share of financial costs in the margin structure (CF/NS), and Vego Supermercados (VS) the largest depreciation and amortization over fixed assets (AAD/FA).

Although formal CoDa methods for clustering firms according to financial statement structure have been developed (Linares-Mustarós et al., 2018), the biplot makes for an appealing visual classification of strategic groups. For instance, firms close to the origin of the rays would fit the so-called stuck-in-the-middle strategy (e.g., AlCampo –AC–, Supermercados Sabeco –SS– and Juan Fornes –JF). Alimerka (AK), Mercadona (MD), Vego Supermercados (VS) and Carrefour (CF) seem to form a strategic group characterized by low fixed asset per employee and low share of financial costs in the margin structure. Supercor (SC) and HiperCor (HC) are characterized by low fixed asset turnover and large overhead share in the margin structure, and so on.

Figure 5 shows the covariance biplot with evolution over time (only 2008 and 2015 data shown). In spite of the fact that all observations have been used for its computation, only 2008 and 2015 values are shown for the sake of simplicity. There are two dominant trajectory patterns. Cecosa Supermercados (CE), Carrefour (CF), Mercadona (MD), Supeco Maxor (SM), Supercor (SC) and Vego Supermercados (VS) are examples of a westbound trajectory, depicted by a blue solid arrow. These chains moved towards having lower leverage and lower financial cost share in their cost structure. Ahorramás (AH), Lidl (LD), Condis (CD), Distribuciones Froiz (DF) and AlCampo (AC) are examples of a northbound trajectory shown in a black arrow, whereby chains increase the share of fixed assets in the asset structure and the share of overhead costs in the cost structure. Some chains, namely Alimerka (AK), Dia (DI) and HiperCor (HC), have trajectories in the opposite direction, depicted by broken arrows.

6. Discussion.

This work has presented a data analysis and visualization tool for supporting strategic management decisions. This approach uses the same type of financial statement information used by financial ratios but does not depend on the often arbitrary decision about which financial ratios to compute. Instead of drawing from any particular ratio, the approach maps accounts and other management magnitudes in a two-dimensional plot, on which firms can be ordered with respect to the ratios of any two magnitudes. Thus, the analysis supplies a global perspective on the financial and economic situation of firms, which is more enriching than that offered by a collection of ratios considered in an isolated way. In this way, the global characteristics of firms can be readily visualized. This tool is especially valuable in the field of strategic management. Analytical techniques such as SWOT analysis or strategic groups are applied more rigorously when based on an adequate metric. Moreover, even the use of conceptual paradigms such as the theory of resource-based analysis (Grant, 1991) or Michael Porter's analysis of industry structure and competitive positioning (Porter, 1980) is enhanced when a metric analysis can be performed. Firms with similar financial patterns and performance can be easily identified, as well as the key ratios according to which they behave similarly or differently. Thus, strategic groups can be considered in terms of the relative position of firms. Particular firms are represented and differences between them are derived by taking into account the multiple dimensions represented by each financial magnitude and all pairwise ratios.

Figure 5. Covariance biplot with evolution over time.
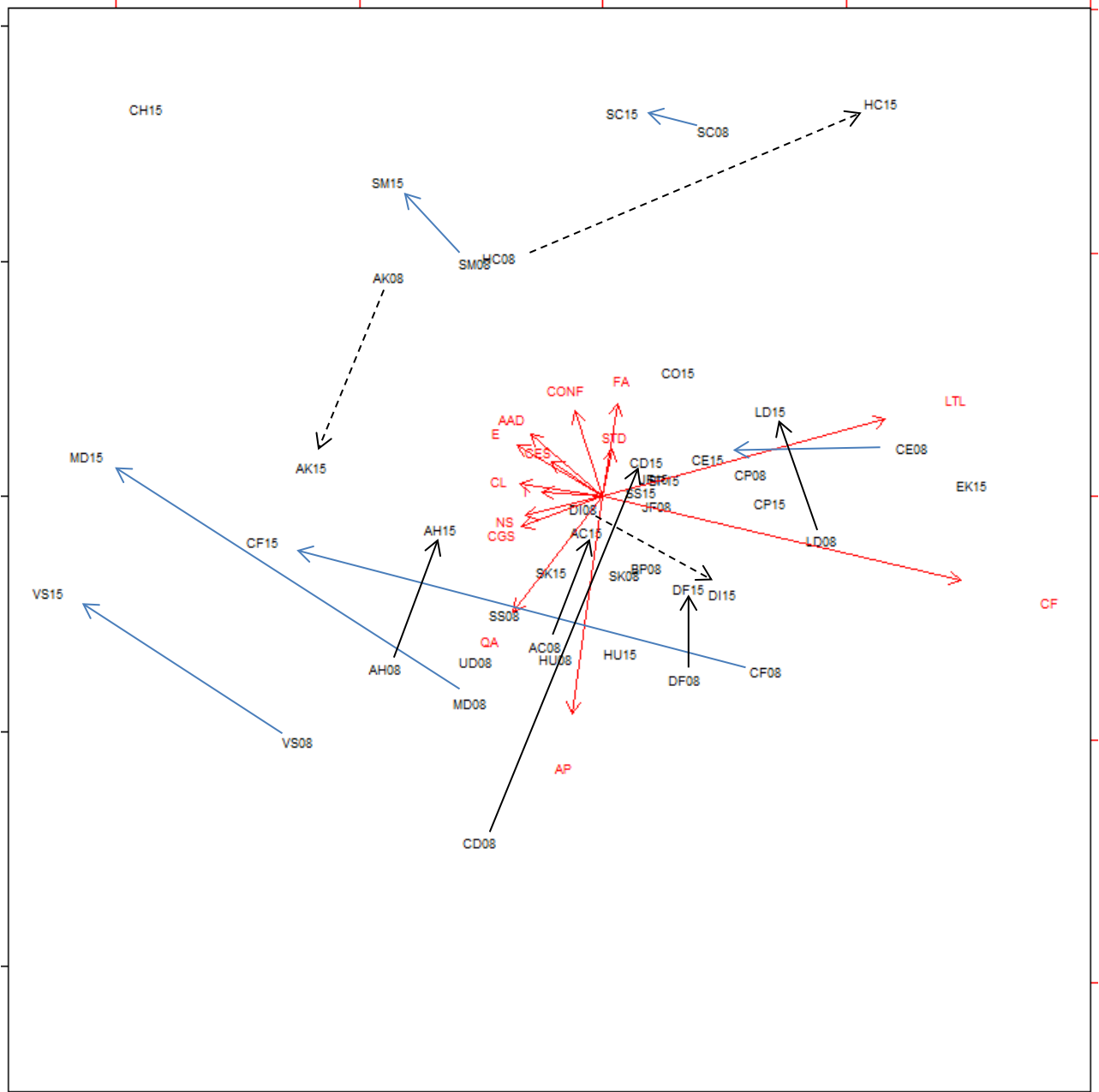
Ahorramás (AH), AlCampo (AC), Alimerka (AK), BonPreu (BP), Caprabo (CP), Carrefour (CF), Cecosa Supermercados (CE), Condis (CD), Consum (CO, outlier in 2008), Dia (DI), Distribuciones Froiz (DF), Eroski (EK, outlier in 2008), HiperCor (HC), Hiper Usera (HU), JuanFornes (JF), Lidl (LD), Mercadona (MD), Semark (SK), Supermercados Champion (CH, outlier in 2008), Supercor (SC), Supeco Maxor (SM), Superercados Sabeco (SS), Unión Detallistas Españoles (UD, outlier in 2015), Vego Supermercados (VS).

Source: Own elaboration.

This tool visualizes the strengths and weaknesses of each firm in a relative way compared to competitors. Moreover, it helps to identify resource-based position by firm. In a global perspective, this tool is useful for describing industrial structure and the competitive positioning of each firm. Therefore, it can be a powerful methodology for assisting in internal strategic analyses. It can also help define the strategic objectives of the company and the subsequent formulation of strategies to be executed, as well as assessing the performance obtained from their implementation. If panel data are available, the individual time evolution of each firm can also be plotted and interpreted in terms of key ratios.

This method draws on the CoDa tradition, and is based on logarithms of ratios, which have proven to lead lo less asymmetry and redundancy than standard financial ratios (Linares-Mustarós et al., 2018)

and not to rely on any particular choice of financial ratios. Another advantage of CoDa in financial analysis is the existence of a formal and well proven method for imputing zero values in financial statements. This makes it possible to perform financial analysis of firms with some zero account values, for which standard financial ratio analysis has to date only provided ad-hoc solutions or has merely dropped them from the analyses.

Several extensions are possible. In this article, mostly financial information has been used to develop and show the proposed methodological approach. However, this tool can deal with any type of quantitative variables, be they financial or not. For example, in the analysis of large distribution chains, we might have considered aspects as relevant as the number of establishments or the commercial area of each retail chain, the volume of Internet sales or other relevant data. All this information can be analysed using this methodological tool if it is in a quantitative metric and non-negative. The method can also be used as a data reduction tool and the first few coordinates explaining most of the variance can be used as variables in further statistical analyses, as explanatory variables in models predicting bankruptcy or belonging to previously known strategic groups, for example. The analysis can also be extended by using weights (Greenacre, 2018) or robust methods (Templ et al., 2011).

Funding

References

Aitchison, J. (1982). The statistical analysis of compositional data. *Journal of the Royal Statistical Society: Series B (Methodological), 44*, 139-177. DOI: 10.1111/j.2517-6161.1982.tb01195.x.

Aitchison, J. (1983). Principal component analysis of compositional data. *Biometrika*, *70*, 57-65. DOI:10.1093/biomet/70.1.57.

Aitchison, J. (1986). *The Statistical Analysis of Compositional Data. Monographs on Statistics and Applied Probability*. London: Chapman and Hall.

Aitchison, J., & Greenacre, M. (2002). Biplots of compositional data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *51*, 375-392. DOI:10.1111/1467-9876.00275.

Allen, R.S., & Helms, M.M. (2006). Linking strategic practices and organizational performance to Porter's generic strategies. *Business Process Management Journal, 12*, 433-455. DOI: 10.1108/14637150610678069.

Azevedo, L., Daunis-i-Estadella, J., Mateu-Flgueras, G., & Thió-Henestrosa, S. (2011). Flying in compositional morphospaces: evolution of limb proportions in flying vertebrates. *Compositional Data Analysis. Theory and Applications* (V. Pawlowsky-Glahn & A. Buccianti, eds). New York, NY: Wiley. pp. 235-254.

Banker, R.D., Mashruwala, R., & Tripathy, A. (2014). Does a differentiation strategy lead to more sustainable financial performance than a cost leadership strategy? *Management Decision, 52*, 872-896. DOI: 10.1108/MD-05-2013-0282.

Barceló-Vidal, C., & Martín-Fernández, J.A. (2016). The mathematics of compositional analysis. *Austrian Journal of Statistics*, *45*, 57-71. DOI: 10.17713/ajs.v45i4.142.

Barnes, P. (1987). The analysis and use of financial ratios: A review article. *Journal of Business Finance & Accounting*, *14*, 449-461. DOI: 10.1111/j.1468-5957.1987.tb00106.x.

Batista-Foguet, J.M., Ferrer-Rosell, B., Serlavós, R., Coenders, G., & Boyatzis, R.E. (2015). An alternative approach to analyze ipsative data. Revisiting Experiential Learning Theory. *Frontiers in Psychology, 6,* 1742. DOI: 10.3389/fpsyg.2015.01742.

Belles-Sampera, J., Guillen, M., & Santolino, M. (2016) Compositional methods applied to capital allocation problems. *Journal of Risk, 19*, 1-15. DOI: 10.21314/JOR.2016.345.

Blanco-Oliver, A.J., Irimia-Diéguez, A.I., & Vázquez-Cueto, M.J. (2016) Design of a specific model for predicting micro-entities failure. *Journal of Quantitative Methods for Economics and Business Administration*, *22*, 3-18. https://www.upo.es/revistas/index.php/RevMetCuant/article/view/2336

Boonen, T., Guillén, M., & Santolino, M. (2019). Forecasting compositional risk allocations. *Insurance: Mathematics and Economics, 84*, 79-86. DOI: 10.1016/j.insmatheco.2018.10.002.

Buccianti, A., Mateu-Figueras, G., & Pawlowsky-Glahn, V. (2006) *Compositional Data Analysis in the Geosciences: From Theory to Practice*. London: Geological Society.

Caro, N.P., Arias, V., & Ortiz, P. (2017). Prediction of failure in Latin-American companies using the nearest-neighbor method to predict random effects in mixed models. *Journal of Quantitative Methods for Economics and Business Administration, 24*, 5-24. https://www.upo.es/revistas/index.php/RevMetCuant/article/view/2878.

Chen, K.H., & Shimerda, T.A. (1981). An empirical analysis of useful financial ratios. *Financial Management, 10*, 51-60. DOI: 10.2307/3665113.

Cowen, S.S., & Hoffer, J.A. (1982). Usefulness of financial ratios in a single industry. *Journal of Business Research, 10*, 103-118. DOI: 10.1016/0148-2963(82)90020-0.

Davis, B.C., Hmieleski, K.M., Webb, J.W., & Coombs, J.E. (2017). Funders' positive affective reactions to entrepreneurs' crowdfunding pitches: The influence of perceived product creativity and entrepreneurial passion. *Journal of Business Venturing, 32*, 90-106. DOI:10.1016/j.jbusvent.2016.10.006.

Deakin, E.B. (1976). Distributions of financial accounting ratios: some empirical evidence. *The Accounting Review*, *51*, 90-96.

Dimitropoulos, P.E., Asteriou, D., & Koumanakos, E. (2010). The relevance of earnings and cash flows in a heavily regulated industry: Evidence from the Greek banking sector. *Advances in Accounting, 26*, 290-303. DOI: 10.1016/j.adiac.2010.08.005.

Egozcue, J.J. Pawlowsky-Glahn, V. Mateu-Figueras, G., & Barceló-Vidal, C. (2003). Isometric logratio transformations for compositional data analysis. *Mathematical Geology, 35*, 279-300. DOI: 10.1023/A:1023818214614.

European Union (2016). *Competition in the Food Retail Sector Proceedings of the Workshop*. Policy Department A: Economic and Scientific Policy. www.europarl.europa.eu/supporting-analyses.

Evans, J. R., & Mathur, A. (2014). Retailing and the period leading up to the Great Recession: a model and a 25-year financial ratio analysis of US retailing. *The International Review of Retail, Distribution and Consumer Research, 24*, 30-51. DOI:10.1080/09593969.2013.801360.

Ferrer-Rosell, B., & Coenders, G. (2018). Destinations and crisis. Profiling tourists' budget share from 2006 to 2012. *Journal of Destination Marketing & Management, 7*, 26-35. DOI:10.1016/j.jdmm.2016.07.002.

Ferrer-Rosell, B., Coenders, G., & Martínez-Garcia, E. (2015). Determinants in tourist expenditure composition-the role of airline types. *Tourism Economics, 21*, 9-32. DOI:10.5367/te.2014.0434.

Ferrer-Rosell, B., Coenders, G., & Martínez-Garcia, E. (2016a). Segmentation by tourist expenditure composition. An approach with compositional data analysis and latent classes. *Tourism Analysis, 21*, 589-602. DOI: 10.3727/108354216X14713487283075.

Ferrer-Rosell, B., Coenders, G., Mateu-Figueras, G., & Pawlowsky-Glahn, V. (2016b). Understanding low cost airline users' expenditure pattern and volume. *Tourism Economics, 22*, 269-291. DOI:10.5367/te.2016.0548.

Filzmoser, P., Hron, K. & Templ, M. (2018). *Applied Compositional Data Analysis with Worked Examples in R*. New York: Springer.

Frecka, T.J., & Hopwood, W.S. (1983). The effects of outliers on the cross-sectional distributional properties of financial ratios. *The Accounting Review, 58*, 115-128.

Gabriel, K.R. (1971). The biplot-graphic display of matrices with application to principal component analysis. *Biometrika, 58*, 453-467. DOI:10.1093/biomet/58.3.453.

Glassman, D.A., & Riddick, L.A. (1996). Why empirical international portfolio models fail: evidence that model misspecification creates home asset bias. *Journal of International Money and Finance, 15*, 275-312. DOI: 10.1016/0261-5606(95)00046-1.

Grant, R.M. (1991). The resource-based theory of competitive advantage: implications for strategy formulation. *California Management Review, 33*, 114-135. DOI: 10.2307/41166664.

Grant, R.M. (2008). *Contemporary Strategy Analysis*. Malden, MA: Blackwell Publishing.

Greenacre, M. (2018). *Compositional Data Analysis in Practice*. Boca Raton, FL: CRC Press.

Hoque, Z. (2004). A contingency model of the association between strategy, environmental uncertainty and performance measurement: impact on organizational performance. *International Business Review, 13*, 485-502. DOI:10.1016/j.ibusrev.2004.04.003.

Horrigan, J.O. (1968). A short history of financial ratio analysis. *The Accounting Review, 43*, 284-294.

Joueid, A., & Coenders, G. (2018). Marketing innovation and new product portfolios. A compositional approach. *Journal of Open Innovation: Technology, Market and Complexity, 4,* 19. DOI: 10.3390/joitmc4020019.

Linares-Mustarós, S., Coenders, G., & Vives-Mestres, M. (2018). Financial performance and distress profiles. From classification according to financial ratios to compositional classification. *Advances in Accounting, 40*, 1-10. DOI: 10.1016/j.adiac.2017.10.003.

Lovell, D., Pawlowsky-Glahn, V., Egozcue, J.J., Marguerat, S., & Bähler, J. (2015). Proportionality: a valid alternative to correlation for relative data. *PLoS Computational Biology, 11*, e1004075. DOI: 10.1371/journal.pcbi.1004075.

Lukason, O., & Laitinen, E.K. (2016) Failure processes of old manufacturing firms in different European countries. *Investment Management and Financial Innovations, 13*, 310-321.

Mariné-Roig, E., & Ferrer-Rosell, B (2018). Measuring the gap between projected and perceived destination images of Catalonia using compositional analysis. *Tourism Management, 68*, 236-249. DOI: 10.1016/j.tourman.2018.03.020.

Martín-Fernández, J.A., Palarea-Albaladejo, J., & Olea, R.A. (2011). Dealing with zeros. *Compositional Data Analysis. Theory and Applications* (V. Pawlowsky-Glahn & A. Buccianti, eds). New York, NY: Wiley. pp. 47-62.

Martín-Oliver, A., Ruano, S., & Salas-Fumás, V. (2017). The fall of Spanish cajas: Lessons of ownership and governance for banks. *Journal of Financial Stability, 33*, 244-260. DOI:10.1016/j.jfs.2017.02.004.

Mateu-Figueras, G., Daunis-i-Estadella, J., Coenders, G., Ferrer-Rosell, B., Serlavós, R., & Batista-Foguet, J.M. (2016). Exploring the relationship between two compositions using canonical correlation analysis. *Metodološki Zvezki, 13*, 131-150. http://ibmi.mf.uni-lj.si/mz/2016/no-2/p5MZ13_2.pdf

McGee, J., & Thomas, H. (1986) Strategic groups: theory, research and taxonomy. *Strategic Management Journal, 7*, 141-160. DOI: 10.1002/smj.4250070204.

Morais, J., Thomas-Agnan, C., & Simioni, M. (2018). Using compositional and Dirichlet models for market-share regression. *Journal of Applied Statistics, 45*, 1670-1689. DOI: 10.1080/02664763.2017.1389864.

Norman, P.M. (2018). An exercise to integrate strategic and financial analysis. *Management Teaching Review, 3*, 252-264. DOI: 10.1177/2379298117752680.

OECD (2013). *Competition Issues in the Food Chain Industry*. https://www.oecd.org/daf/competition/CompetitionIssuesintheFoodChainIndustry.pdf

Ortells, R., Egozcue, J.J., Ortego, M.I., & Garola, A. (2016). Relationship between popularity of key words in the Google browser and the evolution of worldwide financial indices. *Compositional Data Analysis. Springer Proceedings in Mathematics & Statistics, Vol. 187* (J.A. Martín-Fernández & S. Thió-Henestrosa, eds). Cham, CH: Springer. pp. 145-166.

Pablos, R.R., Vara, A.J., & Roche, I.C. (2013). Estrategias de las principales empresas de distribución minorista de gran consumo ante la crisis económica. *Distribución y Consumo, 23*, 5-17.

Palarea-Albaladejo, J., & Martín-Fernández, J.A. (2008). A modified EM alr-algorithm for replacing rounded zeros in compositional data sets. *Computers & Geosciences, 34*, 902-917. DOI: 10.1016/j.cageo.2007.09.015.

Palarea-Albaladejo, J., & Martín-Fernández, J.A. (2015) zCompositions-R package for multivariate imputation of left-censored data under a compositional approach. *Chemometrics and Intelligent Laboratory Systems, 143*, 85-96. DOI: j.chemolab.2015.02.019.

Pawlowsky-Glahn, V. & Buccianti, A. (eds) (2011). *Compositional Data Analysis. Theory and Applications*. New York, NY: Wiley.

Pawlowsky-Glahn, V., Egozcue, J.J., & Tolosana-Delgado, R. (2015). *Modeling and Analysis of Compositional Data*. Chichester, UK: Wiley.

Pinches, G.E., Mingo, K.A., & Caruthers, J.K. (1973). The stability of financial patterns in industrial organizations. *The Journal of Finance, 28*, 389-396. DOI: 10.1111/j.1540-6261.1973.tb01782.x.

Porter, M. (1980). *Competitive Strategy: Techniques for Analysing Industries and Competitors*. New York, NY: The Free Press.

Ross, S.A., Westfield, R.W., & Jordan, B.D. (2003). *Fundamentals of Corporate Finance*. 6th ed. (Vol.1). New York, NY: McGraw-Hill.

Sanz, J.Á., Bedate, A.M., & Durántez, M. (2018). Determining some factors of the financial situation in the European Union publishing sector. *Review of Economic Perspectives, 18*, 25-43. DOI: 10.2478/revecp-2018-0002.

Sharma, S., Shebalkov, M., & Yukhanaev, A. (2016). Evaluating banks performance using key financial indicators-a quantitative modeling of Russian banks. *The Journal of Developing Areas, 50*, 425-453. DOI: 10.1353/jda.2016.0015.

Smith, M. (2005). *Performance Measurement & Management: a Strategic Approach to Management Accounting*. London: SAGE.

Sudarsanam, P.S., & Taffler, R.J. (1995). Financial ratio proportionality and inter-temporal stability: An empirical analysis. *Journal of Banking & Finance, 19*, 45-60. DOI: 10.1016/0378-4266(94)00044-4.

Templ, M., Hron, K., & Filzmoser P. (2011). robCompositions: an R-package for robust statistical analysis of compositional data. *Compositional Data Analysis. Theory and Applications* (V. Pawlowsky-Glahn & A. Buccianti, eds). New York, NY: Wiley. pp. 341-355.

Thió-Henestrosa, S., & Martín-Fernández, J.A. (2005). Dealing with compositional data: The freeware CoDaPack. *Mathematical Geology, 37*, 773-793. DOI: 10.1007/s11004-005-7379-3.

Van den Boogaart, K.G., & Tolosana-Delgado, R. (2013). *Analyzing Compositional Data with R*. Berlin: Springer.

Van Eijnatten, F.M., van der Ark, L.A., & Holloway, S.S. (2015). Ipsative measurement and the analysis of organizational values: an alternative approach for data analysis. *Quality & Quantity, 49*, 559-579. DOI: 10.1007/s11135-014-0009-8.

Vives-Mestres, M., Daunis-i-Estadella, J., & Martín-Fernández, J.A. (2014). Out-of-control signals in three-part compositional T^2 control chart. *Quality and Reliability Engineering International, 30*, 337-346. DOI: 10.1002/qre.1583.

Vives-Mestres, M., Martín-Fernández, J.A., & Kenett, R. (2016a). Compositional data methods in customer survey analysis. *Quality and Reliability Engineering International, 32*, 2115-2125. DOI: 10.1002/qre.2029.

Vives-Mestres, M., Daunis-i-Estadella, J., & Martín-Fernández, J.A. (2016b). Signal interpretation in Hotelling's T^2 control chart for compositional data. *IIE Transactions, 48*, 661-672. DOI: 10.1080/0740817X.2015.1125042.

Voltes-Dorta, A., Jiménez, J.L., & Suárez-Alemán, A. (2014). An initial investigation into the impact of tourism on local budgets: A comparative analysis of Spanish municipalities. *Tourism Management, 45*, 124-133. DOI: 10.1016/j.tourman.2014.02.016

Wang, H., Lu, S., & Zhao, J. (2019). Aggregating multiple types of complex data in stock market prediction: A model-independent framework. *Knowledge-Based Systems, 164*, 193-204. DOI: 10.1016/j.knosys.2018.10.035.

Yap, B.C.F., Mohamed, Z., & Chong, K.R. (2014). The effects of the financial crisis on the financial performance of Malaysian companies. *Asian Journal of Finance & Accounting, 6*, 236-248. DOI: 10.5296/ajfa.v6i1.5314.

Yoshino, N., & Taghizadeh-Hesary, F. (2015). Analysis of credit ratings for small and medium-sized enterprises: Evidence from Asia. *Asian Development Review, 32*, 18-37. DOI: 10.1162/ADEV_a_00050.

Yoshino, N., Taghizadeh-Hesary, F., Charoensivakorn, P., & Niraula, B. (2016). Small and medium-sized enterprise (SME) credit risk analysis using bank-lending data: An analysis of Thai SMEs. *Journal of Comparative Asian Development, 15*, 383-406. DOI: 10.1080/15339114.2016.1233821.