Auditing Issues across Countries: an Explorative Approach to the Regulation Framework

D'Agosto, Elena

May 2007

Auditing Issues across Countries: an Explorative Approach to the Regulation Framework*

Elena D'Agosto

May 28, 2007

Abstract

This paper presents findings from a study of the auditing characteristics in some countries. Auditing rule differentiation among countries was investigated through an empirical analysis based on a multivariate methodological approach. Differences were found in terms of regulation guarantees and the prevention of agency problems.

Key words: Correspondance analysis; Auditing.
JEL: C8; M42.

*The author thanks Prof. Michele Bagella, Prof. Leonardo Becchetti and Nazaria Solferino for discussion and comments. The author is grateful to Prof. Roberto Rocci for having assisted in the drawing up of the methodological appendix. The author also thanks Sabrina Auci to have shared data collection efforts. Any defects in this paper remain the sole responsibility of the author.

1 Introduction

In the last decade some major auditing companies have been charged with negligence in the execution of their duties. Scandals involving these firms brought to light the fact that the main accounting control systems were ineffective.

The theoretical and empirical literature of accounting and business have emphasized different issues associated with the auditing activities and specifically conflicts of interest, fee dependence and mandatory rotations. These are crucial aspects affecting the corporate governance system in a country and the protection of investors' interests.

In this paper we consider how legislation in different countries has tackled auditing issues by looking through audit regulations in the United States (henceforth, the US), the United Kingdom (henceforth, the UK) and four other European countries: France, Germany, Italy and Spain. The idea is to create a quality measure which captures the audit characteristics in the countries inspected. To this end, two main steps are required. The first one is to set up a database compiling several features of auditing for the six countries; the second one is to apply a suitable methodology to our qualitative data to obtain an audit measure. Prior to the data collection we require to identify how auditing and its actors have been ruled by different legislation. So, we sought legal sources for each country investigated and then listed a set of characteristics suitable to represent the auditing rule framework operating in a country. We based the selection of attributes on the Sarbanes Oxley Act (2002). This Act was issued by the US Parliament in the wake of numerous US corporate failures and estabilished standards for all US public company boards, management and accounting firms. Using this law, we consider several main auditing characteristics to compare the audit regulations in the other countries. Further we have checked these features over a period of twenty years in order to capture the changes in the regulation. The list of attributes has been grouped into three clusters denoted as: Supervisory Authority, Auditors' Independence and Corporate Responsibility. The first group lists key issues concerning the authority that oversees the auditing firms' activities. The second cluster describes several aspects relating to Auditors' Independence and the last group illustrates aspects of Corporate Responsability in the auditing process. For each feature we provide a short description and legal sources. Next we explore the data collection. By means of a multiple correspondence analysis, we describe and investigate the evolution in the auditing legislation and make some comparisons between them. This statistical technique allows the reduction of the listed characteristics into a small number of attributes (latent factors). We illustrate the basic features of this approach and then implement it with reference to the auditing dataset. We then present the main results of the analysis.

This paper proceeds as follows. Section 2 mentions some theoretical questions of auditing activity. Section 3 describes the auditing clusters and the conversion to qualitative data. In Section 4 we briefly address the methodology adopted to analyse the data. Section 5 illustrates the main results and breaks these down for each country. The last section contains a brief review of the analysis.

2 Some theoretical aspects of auditing activity

An important outcome of auditing is to make corporate information available to the financial markets through the auditors' statements. Auditors verify and testify to the truthfulness and correctness of financial and economic reports by issuing a certificate. Their aim is to certify (to declare) the validity and the accuracy of a company's reports. This provides signals to market agents on the reliability of results attained by firms, anabling them to make investment decision with confidence.

Historically, the introduction of auditors for checking accounting mistakes or irregularities was due to the development of separation of ownership and control in companies during the 19th century. The agency theory underlines the fact that the main concern stemmed from a possible conflict between managers and shareholders: the audit is a mechanism introduced to minimize managers' incentives to act in their own interests rather than those of shareholders. As a result, the role of the auditing is to reduce contractual or transaction costs related to asymmetric information among parties of an obligation (Jensen and Meckling, 1976). Several empirical studies have verified this link and they have shown a positive relationship between variables representing agency costs and variables expressing audit qualification (Chow, 1982). It has also been shown that agency costs have a significant role in the corporate decision process on auditor selection (Francis and Wilson, 1988 and Defond, 1992).

Auditing quality depends on the ability of the auditor to carry out a careful examination of accounts and to detect possible errors or anomalies (technical competence) and his keenness to give an objective opinion on them (his independence). Although this last factor is a crucial aspect, it is also difficult to observe and to measure. A broad range of theoretical and empirical literature has examined the nexus between independent judgement and auditing provisions, defining this link in terms of fee dependence. It consists of the fee effect (fee provided by customers for audit services) on the auditor independence judgement. The results of these studies are not consistent. For instance, Craswell, Francis and Taylor (1995) affirm that auditing firms, operating in competitive markets, provide lower audit independence judgement if they get provisions which can be replaced. In Wines (1994), Barkess and Simnett, (1994) and Craswell and Francis (1999), assuming the hypothesis of fee dependence, they conclude that this factor does not affect auditor judgments on balance sheets. Moreover, Craswell and Francis (1999) and Craswell et al. (1995) examine the importance of fee dependence, considering audit firms operating at both an international and local level. Their findings do not conclude that fee dependence is a threat to auditor independence.

Another strand of literature studying audit quality highlights the relationship between auditor accuracy and audit firm size. There are two main approaches. The first one is due to De Angelo (1981). He argues that large auditors have more incentive to avoid inaccurate audit reports in order to maintain their own reputation. In fact, if information spreads that auditing firms issue incorrect reports, they could loose earnings because of a diminishing custmer base[1]. The second approach takes into account the "deep pockets" model. According to this model, large auditors have more incentive to issue correct statements becouse otherwise they may suffer from an increase in the size of the litigation penalty. In particular, Dye (1993) and Schwartz (1997) show a positive relationship between auditor accuracy and the size of litigation penalty in cases of negligent behaviour. Furthermore, Lennox (1999) points out that large auditors not only have more incentive to make the right effort but also that they are more likely to be sued because of audit error accruing.

Recent financial scandals (i.e. Enron, Worldcom) have raised again the debate on the quality of audit services, urging scholars to make additional comments. Some contributions have emphasized the importance of the auditor mandatory rotation.

When audit firms have a long-term relationship with their clients, this fact may be seen as a threat to their independence. The introduction of the mechanism of mandatory rotation has the aim of reducing the possibility that a client may influence the auditor by threatening him with the revoke of the contracts.

Despite this importance, empirical research has found ambiguous results. Arrunada and Paz Ares (1997) have analyzed the effect of mandatory rotation on audit cost and quality. Authors show that rotation increases audit cost and price through the destruction of specific assets (i.e. customer knowledge) and the distortion of competition. Besides, they consider that a highly plausible effect is a negative impact on quality, as a consequence of the lower technical competence of the new auditors and ,at least in some cases, fewer incentives for independent behaviour.

Other analyses stress the effects of rotation on total audit costs. Arrunada (2000) illustrates that mandatory rotation raises costs for both auditor and audited firms. Mandatory rotation leads to an increase of direct implicit and explicit costs relating to initial auditing and it give rise to an indirect effect in audit efficiency because of less competitive markets. Finally, the author shows that large audit firms flatten their market share reducing competitiveness, along with the increasing of rotation.

Recently the debate on audit quality and independence has focused on the effects of non-audit services provided by audit firms in the audit market. Audit firms provide complementary services (*non-audit services*) which benefit fom economies of scope. These are due to the provision of both the main and complementary services together, along with a lower cost for providing each services. However, empirical literature does not provide consistent results about the effects of supplying non-audit services on auditors' independence (Craswell, 1999; Parkash and Venable, 1993; Arrunada, 1999). The reasons for this are based on the auditors quasi-rent, in contrast to both the specific and global quasi- rent for client.

[1] Existing customers may not renew contracts and/or a lack of new customers.

This concern arose crucially with the Enron affair and similar cases where auditing firms were involved in the financials scandals; this then created the basis for the Sarbanes Oxley Act in 2002.

3 Auditing Database

In this section, we describe the information gathering process we carried out to create the audit database. We have consulted different legal sources for each country included in the sample. We look at laws pertaining to auditing rules and specifically at company laws. Company laws are part of the commercial codes in civil law countries (France, Spain, Italy and Germany) or exist as separate laws, mainly in the form of acts, in common law countries (US, UK).

In Table 1 we show the legal sources from which the auditing rules have been acquired including Commercial Codes, Company Acts and Codes of Best Practise. Prior to the data collection we selected a list of features that could capture the quality of auditing rules in a given country. We identified these characteristics following the Sarbanes Oxley Act (SOX) which was passed by the US Parliament in 2002 to overcome auditing and accounting shortcomings arising from the financial scandals.

In the next paragraph we consider in more detail which aspects we took into account to represent the quality of the auditing activity.

3.1 Auditing clusters

In Table 2, we present the list of characteristics selected to focus on the quality of auditing. Each of these attributes can be thought of as a measure of an aspect of auditing. We arrange these features into three clusters corresponding to the basic aspects of auditing.

The first cluster is labelled "**Supervisory Authority**" and brings together aspects describing the actors and tasks of this institution, as follows. *Juridical Nature of Authority* defines the legal position of the authority and establishes its entire independence from other public institutions, within the limits of the law. *Duties of the Board* lists the principal duties of the authority relating to the elimination or reduction of probable conflicts of interests. These include, for example, the duty of holding a public accounting firms' register. The auditors are obliged to be on the register in order to practise. The authority has to establish auditing and quality standards and professional ethics standards. Finally, it lists another duty: the possibility of inspections and investigations on public accounting firms' actions. *Composition* refers to how the law defines the main characteristics that each member must have to participate in the authority board. Essentially, members must be chosen based on good reputation and moral integrity. Furthermore, to respect the independence principle of members, as characterized by *Members Independence*, the law bans members from doing other professional activities. In particular, nobody can hold shares of or be paid by a public accounting firm. The aim of this rule is to avoid any conflicts of interest between inspectors and the companies they are checking. *Vacancies* refers to how the law provides rules to appoint a new authority board. The legislator seeks to eliminate the possibility of a period of authority inactivity. *Powers and Rules of Authority* defines whether powers are well separated from the duties of the authority. Some of the most important of these are the powers to: i) sue and be sued, to complain and to defend itself in judicial court; ii) manage the money of public funds as it sees fit respecting the budgetary limits; iii) accept gifts or donations and iv) enter into contracts, execute instruments, incur liabilities, and do any and all other acts and things necessary to the conduct of its operations. *Registration with the Authority*, the law provides that public accounting firms are obliged to be entered on the public register held by the authority. In this way, the authority can control the independence and qualifications of auditors. *Auditing and Quality Control Standards* imply that the authority has the right to establish both auditing standards and both quality and ethics principles. Public accounting firms have to follow these standards in order to continue to work. *Independence Standards and Rules*: according to this attribute, the authority establishes such rules as may be necessary or appropriate in the public interest or for the protection of investors.

In particular, shareholders protection becomes tougher with the analysis of *Inspections of Registered Public Accounting Firms.* The supervisory authority is able to conduct an inspection identifying any act, practise or omission that is in violation of the law or auditing standards. Moreover, the authority can begin a formal investigation or take appropriate disciplinary action, if necessary, with respect to any violations.

If a public accounting firm is unjustly inspected and is subjected to unfair disciplinary actions, the law provides the possibility to sue another authority in the judical court. This is included in *Authority Review.*

Investigations and Disciplinary Proceedings: the law establishes that the authority may conduct an investigation of any act or practice or omission by a registered public accounting firm that may violate rules or standards relating to the preparation and issuance of audit reports. Moreover, the authority may impose some disciplinary measures or sanctions such as temporary suspension or permanent revocation of registration or limitation on activities, functions, or operations of such firm or person.

Applicability to Foreign Public Accounting Firms. This provides the possibility that national rules may be applied to foreign public accounting firms operating in the country.

Finally, *Oversight of Supervision Authority* establishes that another institution exists with the power of control on audit authority's actions.

The second cluster relates to several aspects of the **auditors' independence**. We consider the features of independence of auditors with respect to certified firms. The first element concerns the *Prohibited Activities.* These summarise all activities that a public accounting firm can or cannot do if it is engaged to perform an audit. Among different prohibited activities, we can consider: i) bookkeeping or other services related to accounting records or financial statements of the audit client; ii) giving financial information or acting as broker or dealer or supplying investment banking services; iii) offering internal audit outsourcing services; iv) submitting legal and expert services unrelated to the audit. *Audit Partner Rotation*: the law provides directly that the appointment of a public accounting firm cannot be renewed an unlimited number of times.

Finally, *Conflicts of Interest* prohibs a registered public accounting firm to perform any audit service, if a chief executive officer, controller, chief financial officer or chief accounting officer is employed by that registered independent public accounting firm and participated in any capacity in the audit of that listed firm.

In the final cluster, referred to as "**Corporate responsibility**", we take into account features related to the responsibilities of internal audit committee and of each member the board of directors. *Audit Committees*: this identifies responsibilities of the internal audit committee. Furthermore, the majority of the committee's members must be independent from the board, in which they can participate on a purely advisory basis.

Corporate Responsibility for Financial Reports of National and Foreign Firms summarizes the responsibilities of the board with respect to financial and economic statements. The law provides that balance sheets have to be accompanied by a written statement by the chief executive officer and the chief financial officer. This report has to certify the truthfulness and completeness of the information in balance sheets.

Improper influence on Conduct of Audits: it is an illegal act for any officer or director of an issuer, to take any action to fraudulently influence, coerce, manipulate, or mislead any independent public accounting firm, which is engaged in the performance of an audit of financial statements for the purpose of rendering such financial statements materially misleading.

CEO and CFO reimburse the Issuer due to non-compliance of financial reports. This considers the possibility that the chief executive officer and the chief financial officer of the issuer have to reimburse the issuer for any bonus or other incentive-based or equity-based compensation received, in the case of their misconduct.

3.2 Measuring audit characteristics

Since our purpose is to compare audit regulations across countries, we first turn each of the audit rules (listed in table 2) into qualitative variables. To this aim, we create variables which assume value 1 if the specific auditing rule is present in the country legislation and zero otherwise. We obtain a matrix of binary variables called an indicator matrix. We repeat this exercise for each year of the sample period from 1980 to 2002, giving a total indicator matrix which consists of 132 rows (22 years times 6 countries) and 22 columns (qualitative variables). The total matrix represents our auditing database. This simple measurement method avoids the needs for a discretionary weighting, such as scores for each country, and allows us to apply a statistical methodology which compares countries with respect to auditing regulations. In the next section we briefly describe this procedure before applying it to our dataset.

4 Data Analysis

In examining our database we expect to capture countries' peculiarities. Multivariate statistical methods provide descriptive and explorative approaches able to inspect data and to aid the identification of a profile of the country audit legislation over time.

Among these methods, correspondence analysis (CA) is accurate for the study of qualitative data and has added features of not assuming an underlying distribution and not hypothesising a model for the data (Greenacre, 1984)[2]. This procedure allows an overview of the fundamental relationships among the variables and removes irrelevant information. Moreover it can display results in graphic form and is easy to interpret.

In the following paragraph we describe the CA method and we refer to the appendix for a more thorough description.

4.1 Methodology

In broad terms, correspondance analysis explores the structure of associations among a set of qualitative variables and brings out underlying dimensions which can be interpreted as latent variables. To illustrate this, let us assume, for instance, that a contingency table has m rows and k columns, each row and each column in the table representing one modality of the corresponding variable. We can assume that the m values of rows in the k^{th} column can be considered as coordinates of the point in an m-dimensional space which summarize all information about the similarities betweeen the rows [3]. The CA technique consists of reducing the dimensional space in which to place points in such a way that most of the information about the differences between the rows is retained. The dimensions of the lower space define the number of axes. Each axis of the new space represents a "synthesis" of several piece of information of the initial space. These axes must be interpreted according to the modalities that have contributed to their definition the most, in order to give a meaning to the graphical solution provided. The amount of information retained in all dimensions of the new space is called total inertia.

In the graphical solution, the locations of all modalities of the variables can be compared to each other: short distances imply high similarities and grater distances high dissimilarities.

A simple correspondence analysis is applied to a single contingency table. When we have a multivariate phenomenon then a multiple correspondence analysis (MCA) is required. This latter technique is in fact the appropriate method in this case.

An appropriate way to treat a multidimentional analysis is to convert data into an indicator matrix. This matrix has each observational unit (in our case,

[2] See also, Greenacre (1995) and Greenacre and Blasius (1995).

[3] In the same and symmetric way we can investigate the similarities between columns. This is the core of the Dual Problem: considering both rows or columns we get analogous results (see appendix for more details).

countries for each year) in the rows and each variable (with its own modalities) located in the columns. Hence, value 1 is assigned to each modality in which a country is placed. It can be showed that Multi-Correspondence analys is the CA procedure applied to an indicator matrix.

In this part of the paper, we summarize conceptually several rules for interpreting results, refering to the appendix for analytical explanation.

By applying the MCA methodology we can obtain some important information.

The first item relates to a quality measure indicating as a percentage, to what extent the chosen dimensions explain the total inertia. In this way we measure the information lost when reducing the dimentional space. It may be interpreted as the percentage of variance explained by the chosen axes.

Another item of information indicates to what extent each variable or individual contributes to the inertia explained by each dimension. This allows us to assess the contribution of each variable or individual to the axis orientation. The co-ordinates of variables or individuals are also considered with respect to the dimensions.

5 Description of the results

The results of the multi-correspondence analysis are presented in two steps. The first one consists of interpreting the extracted axes. Then, we consider the position of the countries within the reduced dimensional space.

Primarily, we introduce the general outcomes of the analysis which are reported in table 3. This shows the decomposition of inertia with respect to the principal axes and contains the eigenvalues of the matrix. Each axis accounts for a part of the inertia, and it is also expressed as a percentage.

Figure 1 displays the histogram of the inertia expressed by each axis which may be interpreted as its own percentage of explained variance. The outcomes show that the first principal axis explains about half of the total variance and the two first axes together express two thirds of the total variance. The first dimension explains 48.44% of the total inertia. The second one accounts for 16.21% for a total of 64.65%. We consider this cumulative inertia adequate to summarize the information retained from the original space. Therefore we will only interpret these axes. This result is particularly interesting because we considerably reduce the dimensions of the space while preserving most of the relevant information.

Figure 2 presents the first two principal axes produced by an MCA of the data. In this "map" all the countries have been projected according to their co-ordinates on the two first dimensions. All countries are represented here as points in the plane. The co-ordinates of the countries are calculated on the basis of values assegned to the profile of attributes that describe each axis.

The plane has been constructed in such a way that the maximum "difference" or "contrast" has been created, with regard to attribute profiles, between those countries situated on the right-hand side of the plane and those on the left-hand side. This is the first, the horizontal dimension. Similarly, the second largest possible "difference" with regard to attribute profile, is represented by the "difference" or "contrast" between the countries situated in the top part of the plane, and those at the bottom. This second, vertical dimension, represents this difference.

We observe that countries assembled in the same group (points close to each other) are characterized by homogeneity of attributes.

5.1 Interpretation of the principal axes

We shall now examine the characteristics of the extracted two principal axes. The presentation is organized around one graph and tables that contain the basic results for each of the axes.

Axis one

The first column of table 4 shows the modalities of the variables describing the first axis. These modalities come from both the Supervisory and Auditors Independence cluster. Looking at these aspects we can define the first dimension as *Degree of Regulation Guarantees*. The label applied to the axis means that the information it summarizes is related to the definition of rules which specify the supervisory and auditing activities. Indeed some modalities concern rules about standards of auditing and quality control, some are related to powers and duties of the authority and others consider rules about features on auditors' independence with respect to the certified firms. Therefore the position of a country relative to this axis allows an evaluation of whether a country's legislation assures investors' protection through clear rules that regulate contents and actors involved in auditing activities. To understand the information summarized in this first axis, we first decribe the orientatation of the axis. In Figure 2 all positive values identify the absence of the modalities which characterize the dimension, while negative values identify their presence. The first axis is represented by the horizontal line in which groups of countries located in the right-hand part of the first axis show a lower degree of regulation guarantees with respect to those on the left-hand side. Countries that have similar modalities are represented by points close to each other. Moreover a country that has changed its legislation over time in favour of a better regulation is shown on the first axis with points moving from the right to the left-hand side.

Axis two

Let us turn to the interpretation of the second principal axis extracted in the multiple correspondence analysis of the variables. In Figure 2 the second axis is represented by the vertical line and the main results are given in the Table 5. In this axis we find a measure which indicates whether a regulatory environment is more inclined towards the agency issues. The elements that identify this axis mainly concern corporate and auditor responsibility. They are related to rules that incentivize the agents (auditors or firms) to act in the best interest of investors. For this reason we denote the axis as *Degree of prevention of the agency problem.* In Figure 2, groups of countries that are located in the top part of the second axis show a lower degree of prevention of the agency problem with respect to those in the bottom part. Therefore a country that has changed its legislation over time in favour of better prevention rules is shown with points moving from the top to the bottom part.

Country Results

Now we turn to the analysis of the position of countries on the above two dimensions. Primarly, we consider the location on the first axis. Figure 2 shows the position of each country in the period considered and allows us to track its auditing rules history. It is clear from first glance at the graph that there is a marked distinction between two groups of countries. Those placed on the left-hand side on the graph have a tendency to a high degree of regulation guarantees while for those on the right the converse is true.

The outcomes of the analysis reveal that Germany is located on the left-hand side of the axis throughout the period, illustrating that its legislation has taken auditing rules into account as part of the investors' protection arrangement. Its regulation focuses on characteristics concerning the supervision activity of the authority and some elements that intend to assure auditors' independence. From 2001 it also takes into consideration aspects related to the corporate responsibility, especially responsibility for financial reports and improper influence on auditors. Spain has also given attention to the audit regulatory environment, as confirmed by a shifting of its location relative to the axes through the years. It has demonstrated a high degree of regulation of auditing activities concerning the three actors involved, auditors, firms and the supervisory body since 1986. A further improvement was regarding prohibited auditor activities, has been in force since 1998.

The French system is positioned on the right side of the first axis which is characterized by inadequate precence of regulation guarantees, over the observed period. The Supervision activity does not provide sufficient auditing or independence standards, or rules regarding corporate responsibility. The French system appears to be less inclined to revise its regulatory environment; rather, there is currently a greater attention on auditing activity, although an improvement in audit partner rotation has been in place since 1996.

Turning to the Italian location, the chart indicates a jump from the right to the left side on the first axis. This is due to a new arrangement involving auditing activity rules was made law in 1998.

The United States is placed on the left side of the first dimension from the time when the Sarbanes Oxley Act was passed in 2002. Relative to previous years, the law addresses further concerns regarding the supervisory issues of the auditing activities and corporate responsibility for financial reports. The UK is also situated wholly on the right hand side of the first dimension and any development have tended not to make progress regarding regulation guarantees. However, this appears to be consistent with most common law systems.

Countries that contribute more to the second dimension, are those stressing aspects of the audit rules related to ties between auditing and the audited firm. Countries plotted on the top part of the diagram are differentiated from the others in terms of limited prevention of the agency problems. Interestingly, Italy and Germany contribute to the formation of the axis, characterizing it as having almost no regulation regarding the degree of prevention of the problem of conflict of interest. Moreover, the UK plays a significant role in the formation of the bottom part of the second dimension.

Describing the judicial aspects of these changes is not the aim of this article. However, we report the list of the main laws and regulations that have been passed by each country in the Table 3. This helps to illustrate which changes in national legislation relate to the above descriptions.

6 Conclusion

The aim of this paper is to explore by means of the MCA the evolution of auditing legislation across countries.

To this end, we set up a unique dataset looking through national auditing rules over two decades. The main results of the analysis allow us to define two matters of interest focused on national legislation. The first one is the Degree of regulation guarantee represented by the primary dimension extracted by the MCA. The next is the Degree of prevention of agency problems represented by the second axis.

Looking at the position of the countries on these axes, we can basically divided them into two groups. We observe that for countries such as Italy, Spain and the USA, the attention to regulation guarantee has shifted considerably over time, in terms of auditing framework. This is in contrast to the trends in France and the UK.

Prevention of agency problem is an issue where the UK differs significantly over time compared with France, Italy and Germany.

References

[1] Arruñada B. (1999), "The Provision of Non-Audit Services By Auditors: Let the Market Evolve And Decide" International Review of Law and Economics, 19 (4): 513-31.

[2] Arruñada, B. (1999), "The Economics Of Audit Quality: Private Incentives And The Regulation Of Audit And Non- Audit Services", Dordrecht and Boston: Kluwer.

[3] Arruñada B. (2000), "Audit Quality: Attributes, Private Safeguards and the Role Of Regulation" European Accounting Review, 9(2): 205-224.

[4] Arruñada, B. and Paz-ares C. (1997), "Mandatory rotation of company auditors: a critical examination" International Review of Law and Economics, 17 (March): 31-61.

[5] Barkess, L. and Simnett, R. (1994), "The Provision Of Other Services By Auditors: Independence And Pricing Issues" Accounting and Business Research, 24: 99-108.

[6] Beck, P. J., Frecka T. J. and Solomon I. (1988), "An Empirical Analysis of the Relationship between Mas Involvement and Auditor Tenure: Implications for Auditor Independence" Journal of Accounting Literature, 7: 65-84.

[7] Benzecri, J.P. (1982), Histoire et préhistoire de l'Analyse des Données. Dunod, Paris, France.

[8] Carrol, J.D. (1968), A generalisation of canonical correlation analysis to three or more sets of variables - 76th Convention American Psychological Association.

[9] Chow, C. W. (1982), "The Demand for External Auditing: Size, Debt and Ownership" The Accounting Review, (April): 272-292.

[10] Craswell, A.T. (1999), "Does the Provision of Non-Audit Services Impair Auditor Independence?" International Journal of Auditing, (3): 29-40.

[11] Craswell, A.T. and Francis, J.R. (1999), "Pricing Initial Audit engagements: A Test of Competing Theories" The Accounting Review, 74 (2): 201-216.

[12] Craswell, A.T., Francis J.R. and Taylor S.L. (1995) "Auditor Brand Name Reputations and Industry Specializations." Journal of Accounting and Economics, 20: 297- 322.

[13] De Leeuw, J. (1983), On the prehistory of correspondence analysis. Statistica Neerlandica, 37 (1): 61-164.

[14] DeAngelo, H. and DeAngelo L. (1985), "Managerial Ownership of Voting Rights." Journal of Financial Economics, 14: 33-69.

[15] DeAngelo, L. E. (1981), "Auditor Size and Audit Quality" Journal of Accounting and Economics, (3): 183-99.

[16] Defond, M. (1992), "The Association between Changes in Client Firm Agency Costs and Auditor Switching" Auditing: A Journal of Practice and Theory, (11): 16-31.

[17] DeFond, M. L. and Park C. W. (1999), "The Effect of Competition on CEO Turnover" Journal of Accounting and Economics, 27 (February): 35-56.

[18] DeFond, M. L., Francis, J. R. and Wong T. J. (2000), "Auditor Industry Specialization and Market Segmentation: Evidence from Hong Kong" Auditing: a Journal of Practice and Theory, (Spring): 49-66.

[19] DeFond, M. L., Raghunandan, K. and Subramanyam K. R. (2002), "Do Non-Audit Service Fees Impair Auditor Independence? Evidence from Going Concern Audit Opinions" (University of Southern California, January, Manuscript).

[20] Dye, R. A. (1986), "Proprietary and Nonproprietary Disclosures" Journal of Business, 59 (2): 331-366.

[21] Dye, R. A. (1991), "Informationally Motivated Auditor Replacement" Journal of Accounting and Economics, 14: 347-374.

[22] Francis, J. R. and Wilson E. R. (1988), "Auditor Changes: A Joint Test of Theories Relating to Agency Costs and Auditor Differentiation" Accounting Review, 63 (October): 663-82.

[23] Greenacre M. J. (1995): Correspondence Analysis in Practice. London: Academic Press, INC.

[24] Greenacre M. J., Blasius, J. (1995) Correspondence Analysis in the Social Sciences. London: Academic Press, INC

[25] Greenacre, M. J. (1984), Theory and Applications of Correspondence Analysis, Academic Press, London.

[26] Hotelling, H. (1936), Relations between two sets of variables, Biometrika, 28.

[27] Jensen, M. C. (1993), "The Modern Industrial Revolution, Exit, and the Failure of Internal Control Systems" Journal of Finance, 48: 831-880.

[28] Jensen, M. C. and Meckling W. H. (1976), "Theory of firm: managerial behavior, agency costs and ownership structure" Journal of Financial Economics, 3 (4): 305-360.

[29] Lebart, L., Morineau, A. and Piron, M. (2002), "Statistique exploratoire multidimensionnelle" Dunod, Paris 3° édition.

[30] Lennox, C. S. (1999), "Are Large Auditors More Accurate Than Small Auditors?", Accounting and Business Research, 29 (3): 217-228.

[31] Parkash, M. and Venable C.F. (1993), "Auditee Incentives for Auditor Independence: The Case of Nonaudit Services." The Accounting Review, 68(1): 113-133.

[32] Saporta, G. (1990), Probabilités analyse des données et statistique, Editions Technip, Paris.

[33] Schwartz, R. (1997), "Legal Regimes, audit quality and Investment" The Accounting Review, 72 (July): 385-406.

[34] Wines, G. (1994) "Auditor Independence, Audit Qualifications and the Provision of Non- Audit Services: A Note." Accounting and Finance, 34: 75-86.

Table 1 Source of country audit legislation

COUNTRIES	LAWS:
FRANCE	Commercial Code 1807 Le Plan Comptable Général 1983 Legge del 24 luglio 1966 n. 537
SPAIN	Royal Legislative Decree n. 1564 Stock Corporation Law of 22 December 1989 Act on Auditing n. 19 of 12th July 1988 Royal Decree n. 302 Organic Structure of the Accounting and Auditing Institute of 17th March 1989 The Governance Code of Listed Companies of 26 February 1998 Codice volontario di Corporate Governance
GERMANY	German Stock Corporation Law (AktG) 1965 Commercial Code (HGB) from 1897 Law of Regulating the Profession of Auditor (WPO) 1961 German Act of Corporate Control and Transparency (KonTraG) 1998 Codice volontario di Corporate Governance
UNITED KINGDOM	Joint Stock Companies Act 1844 Companies Act 1985 modified in 1989 Companies Directors Disqualifications Act 1986 Code of Best Practise: Hampel Report 1998 Cadbury Code 1992
ITALY	Codice Civile L. 216/1974 Dlgs. 58/1998 (Draghi law) Dlgs. 2002 Codice di autodisciplina delle società quotate
USA	Securities Exchange Act 1934 sec. 3 Sarbanes- Oxley Act 2002

Table 2: List of audit characteristics

SUPERVISORY AUTHORITY INDICATOR	AUDITORS INDEPENDENCE INDICATOR	CORPORATE RESPONSIBILITY INDICATOR
~ Juridical nature of authority	~ Prohibited Activities	~ Audit Committees: responsibility and independence
~ Duties of the Board	~ Audit Partner Rotation	~ Corporate Responsibility for Financial Reports of National and Foreign Firms
~ Composition	~ Conflicts of Interest	~ Improper influence on Conduct of Audits
~ Members Independence		~ CEO and CFO reimburse the Issuer due to non-compliance of financial reports
~ Vacancies		
~ Powers and Rules of Authority		
~ Registration with the Authority		
~ Auditing Standards		
~ Quality Control Standards		
~ Independence Standards and Rules		
~ Inspections of Registered Public Accounting Firms		
~ Authority Review		
~ Investigations and Disciplinary Proceedings		
~ Applicability to Foreign Public Accounting Firms		
~ Oversight of Supervision Authority		

Table 3 Decomposition of Inertia

Number of axes	Eigen values	Percentage	Cumulative percentage
1	0.4844	48.44	48.44
2	0.1621	16.21	64.65
3	0.1407	14.07	78.71
4	0.1039	10.39	89.10
5	0.0532	5.32	94.43
6	0.0258	2.58	97.01
7	0.0107	1.07	98.08
8	0.0073	0.73	98.81
9	0.0050	0.50	99.31
10	0.0030	0.30	99.61
11	0.0020	0.20	99.81
12	0.0011	0.11	99.93
13	0.0007	0.07	100.00
14	0.0000	0.00	100.00
15	0.0000	0.00	100.00
16	0.0000	0.00	100.00
17	0.0000	0.00	100.00
18	0.0000	0.00	100.00
19	0.0000	0.00	100.00
20	0.0000	0.00	100.00
21	0.0000	0.00	100.00

FIGURE 1 HISTOGRAM OF THE INERTIA EXPRESSED BY EACH AXIS

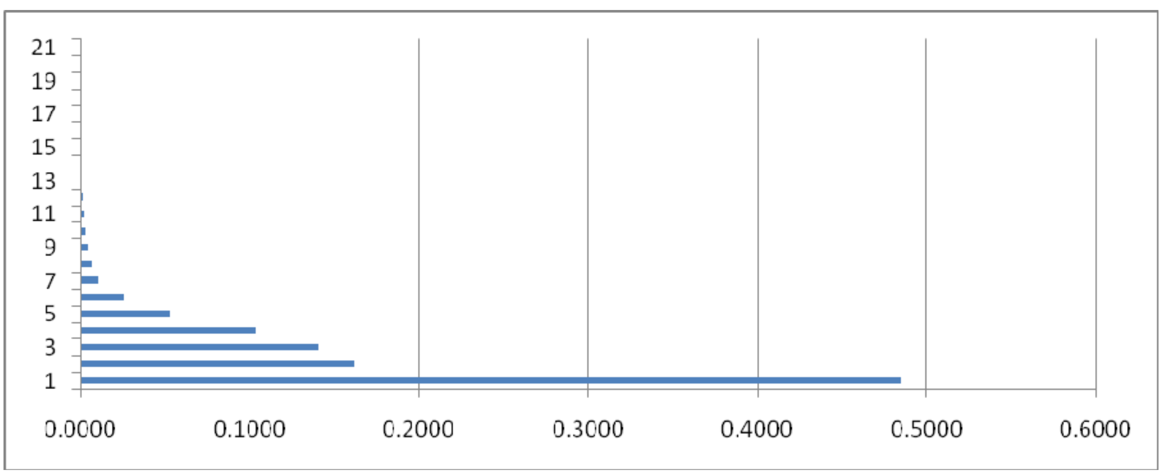

FIGURE 2

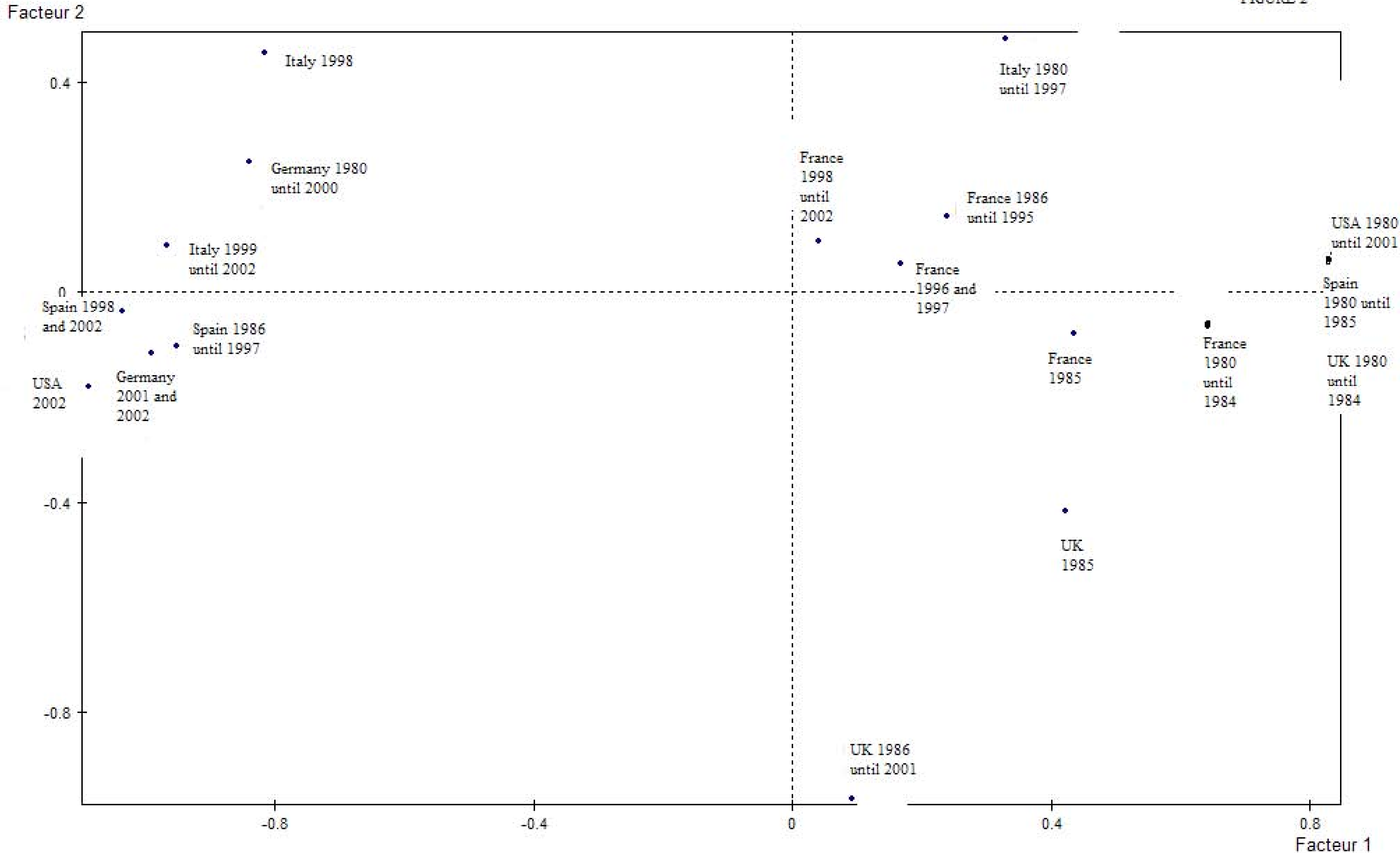
Facteur 2
Facteur 1
0.4
0
-0.4
-0.8
-0.8
-0.4
0
0.4
0.8
Italy 1998
Italy 1980 until 1997
Germany 1980 until 2000
France 1998 until 2002
France 1986 until 1995
Italy 1999 until 2002
USA 1980 until 2001
France 1996 and 1997
Spain 1980 until 1985
Spain 1998 and 2002
Spain 1986 until 1997
France 1980 until 1984
France 1985
UK 1980 until 1984
USA 2002
Germany 2001 and 2002
UK 1985
UK 1986 until 2001

Table 4 Description of the Axis 1

Label of attributes	Modality
Auditing Standards	Present
Independence Standards and Rules	Present
Quality Control Standards	Present
Inspections of Registered Public Accounting Firms	Present
Audit Partner Rotation	Present
Oversight of Supervision Authority	Present
Conflicts of Interest	Present
Registration with the Authority	Present
Duties of the Board	Present
Authority Review	Present
Powers and Rules of Authority	Present
ZONE CENTRALE	
Powers and Rules of Authority	Absent
Authority Review	Absent
Duties of the Board	Absent
Registration with the Authority	Absent
Conflicts of Interest	Absent
Oversight of Supervision Authority	Absent
Audit Partner Rotation	Absent
Independence Standards and Rules	Absent
Quality Control Standards	Absent
Inspections of Registered Public Accounting Firms	Absent
Auditing Standards	Absent

Table 5 Description of the Axis 2

Label of attributes	Modality
Improper influence on Conduct of Audits	Present
Corporate Responsibility for Financial Reports	Present
Audit Committees: responsibility and independence	Present
Composition	Absent
Registration with the Authority	Absent
Independence Members	Absent
Vacancies	Absent
Duties of the Board	Present
Investigations and Disciplinary Proceedings	Present
Conflicts of Interest	Present
Prohibited Activities	Absent
Z O N E C E N T R A L E	
Prohibited Activities	Present
Conflicts of Interest	Absent
Investigations and Disciplinary Proceedings	Absent
Duties of the Board	Absent
Vacancies	Present
Independence Members	Present
Registration with the Authority	Present
Composition	Present
Audit Committees: responsibility and independence	Absent
Corporate Responsibility for Financial Reports	Absent
Improper influence on Conduct of Audits	Absent

Table 6: Values of Axis 1 by country

Years	USA	France	Germany	Italy	Spain	UK
1980	0.84	0.62	-0.85	0.32	0.83	0.83
1981	0.84	0.62	-0.85	0.32	0.83	0.83
1982	0.84	0.62	-0.85	0.32	0.83	0.83
1983	0.84	0.62	-0.85	0.32	0.83	0.83
1984	0.84	0.62	-0.85	0.32	0.83	0.83
1985	0.84	0.43	-0.85	0.32	0.83	0.42
1986	0.84	0.23	-0.85	0.32	-0.96	0.09
1987	0.84	0.43	-0.85	0.32	-0.57	0.42
1988	0.84	0.23	-0.85	0.32	-0.96	0.09
1989	0.84	0.23	-0.85	0.32	-0.96	0.09
1990	0.84	0.23	-0.85	0.32	-0.96	0.24
1991	0.84	0.23	-0.85	0.32	-0.96	0.24
1992	0.84	0.23	-0.85	0.32	-0.96	0.09
1993	0.84	0.23	-0.85	0.32	-0.96	0.09
1994	0.84	0.23	-0.85	0.32	-0.96	0.09
1995	0.84	0.23	-0.85	0.32	-0.96	0.09
1996	0.84	0.16	-0.85	0.32	-0.96	0.09
1997	0.84	0.16	-0.85	0.32	-0.96	0.09
1998	0.84	0.03	-0.85	-0.83	-1.05	0.09
1999	0.84	0.03	-0.85	-0.98	-1.05	0.09
2000	0.84	0.03	-0.85	-0.98	-1.05	0.09
2001	0.84	0.03	-1.01	-0.98	-1.05	0.09
2002	-1.09	0.03	-1.01	-0.98	-1.05	0.09

Table 7: Values of Axis 2 by country

Years	USA	France	Germany	Italy	Spain	UK
1980	0.00	0.11	-0.18	-0.60	0.00	0.00
1981	0.00	0.11	-0.18	-0.60	0.00	0.00
1982	0.00	0.11	-0.18	-0.60	0.00	0.00
1983	0.00	0.11	-0.18	-0.60	0.00	0.00
1984	0.00	0.11	-0.18	-0.60	0.00	0.00
1985	0.00	0.14	-0.18	-0.60	0.00	0.39
1986	0.00	-0.17	-0.18	-0.60	0.15	0.91
1987	0.00	0.14	-0.18	-0.60	-0.26	0.39
1988	0.00	-0.17	-0.18	-0.60	0.15	0.91
1989	0.00	-0.17	-0.18	-0.60	0.15	0.91
1990	0.00	-0.17	-0.18	-0.60	0.15	0.57
1991	0.00	-0.17	-0.18	-0.60	0.15	0.57
1992	0.00	-0.17	-0.18	-0.60	0.15	0.91
1993	0.00	-0.17	-0.18	-0.60	0.15	0.91
1994	0.00	-0.17	-0.18	-0.60	0.15	0.91
1995	0.00	-0.17	-0.18	-0.60	0.15	0.91
1996	0.00	-0.08	-0.18	-0.60	0.15	0.91
1997	0.00	-0.08	-0.18	-0.60	0.15	0.91
1998	0.00	-0.11	-0.18	-0.49	0.09	0.91
1999	0.00	-0.11	-0.18	-0.14	0.09	0.91
2000	0.00	-0.11	-0.18	-0.14	0.09	0.91
2001	0.00	-0.11	0.17	-0.14	0.09	0.91
2002	0.12	-0.11	0.17	-0.14	0.09	0.91

7 APPENDIX

The purpose of this appendix is to present a theoretical review of the technique applied for the derivation of the auditing index. The Multicorrespondence analysis (MCA) is the appropriate method to analyse our dataset.

The starting point should be the examination of the correspondence analysis[4] (henceforth, CA). Since the equations that defines CA can be derived in many different ways[5] (De Leeuw,1983), we decide to explain it in terms of the equivalent approach denoted as canonical correlation analysis (CCA). Then we get the Multiple correspondence analysis as a particular case of the generalized canonical correlation analysis.

The correspondence analysis and the canonical correlation analysis have many parallels with the principal-components analysis (PCA), and so we require to introduce it, shortly.

7.1 AN INTRODUCTION TO THE PRINCIPAL COMPONENTS ANALYSIS

The main idea behind PCA is to replace a set of variables with a smaller number of variables that are not correlated and have maximum variance.

Let $\mathbf{Y}$ denote an $n \times p$ matrix whose element ij is given by y_{ij} and that contains n units and p variables. Let $\mathbf{y}_i = (y_{i1}, y_{i2}, ..., y_{ip},)$ be a vector that consists of p variables associated at unit i, with $i = 1, \ldots\ldots, n$. Let $\tilde{\mathbf{y}}_j = (y_{1j}, y_{2j}, ..., y_{nj},)$ be a vector that consists of n-units associated at variable j. Let us consider that n-units and p-variables are elements of an Euclidean vector space. Then $\mathbf{y}_i$ is a vector defined in a p-dimensional space, $\mathbf{R}^p$ and $\tilde{\mathbf{y}}_j$ is a vector defined in n–dimensional space, $\mathbf{R}^n$.

Let $a_1, ...a_p$ be the canonical basis that define the p-dimensional space. Each vector $\mathbf{y}_i$ can be defined as the linear combination of the basis vectors with coefficients $y_{i1}, y_{i2}, ..., y_{ip},$: $\mathbf{y}_i = y_{i1} a_1 + y_{i2} a_2 + ... + y_{ip} a_p$ Let $\bar{\mathbf{y}}$ be the centroid, or mean vector, of $\mathbf{y}_1 .... \mathbf{y}_n$ defined as the linear combination $\bar{\mathbf{y}} = \alpha_1 y_1 + ... + \alpha_n y_n$ where the coefficients add up to 1: $\sum \alpha_i = 1$.

Let us suppose to represent the cloud of vectors y_i (i=1,. . . . . . ,n) in a lower K^* dimensional subspace $\mathbf{Q}$. Over the sub space, the vectors $\mathbf{y}_i$, $\mathbf{y}_1 .... \mathbf{y}_n$, should maintain as much as possible the informative power of the original space. Geometrically it means that the vectors $\mathbf{y}_1 .... \mathbf{y}_n$ should be projected over the subspace $\mathbf{Q}$ trying to keep the original distance among them. The choices fall over the spaces of lower dimensionality which come closest to the set of original vectors . A way to identify the subspace is to minimize the distance between each vector $\mathbf{y}_i$ and its projection $\hat{\mathbf{y}}_i$ on the K^* dimensional subspace $\mathbf{Q}$.

[4] We refer to the extensive discussion in the books by Greenacre (1984) and by Labart, Morineau and Piron (2000) to write this methodological appendix.

[5] Historical reviews are, for instance, in De Leeuw (1983) and Benzecri (1982). There are two main approaches to the CA: the French approach and the Anglo approach. Respect to the former, initiated by Benzecri around 1965, CA is a form of metric multidimensional scaling. To the second, CA is considered as optimal scoring (scaling) approach.

If $\mathbf{y}_i$ is weighted by a mass w_i then our definition of the closeness of the whole set of vectors in the subspace $\mathbf{Q}$ is: $\Psi(Q;\mathbf{y}_1....\mathbf{y}_n) \equiv \sum w_i d_i^2$ (a.1.1) where $d_i^2 \equiv (\mathbf{y}_i - \hat{\mathbf{y}}_i)' D_s(\mathbf{y}_i - \hat{\mathbf{y}}_i)$ and D_s is the diagonal matrix of the weights assigned to the p original dimensions[6].

The squared distance d_i^2 depends on the subspace $\mathbf{Q}$ and our objective is to find the optimal subspace $\mathbf{Q}*$ which minimizes the function $\Psi(Q;\mathbf{y}_1....\mathbf{y}_n)$.

It can be shown (Greenacre, 1984) that the optimal K^*-dimensional subspace $\mathbf{Q}^*$ is that minimize the function $\Psi(Q;\mathbf{y}_1....\mathbf{y}_n)$ and must contain the mean vector $\bar{\mathbf{y}}$.

The theoretical solution to the problem of minimization for any specified dimensionality K* is embodied in the concepts of singular value decomposition (henceforth denoted by SVD) and low rank matrix approximation.

Let $\mathbf{Y}$ denote an $n \times p$ matrix of rank K, the equation for ordinary SVD of Y is the following:

$\mathbf{Y} = UD_\alpha V'$ with $U\prime U = I = V\prime V$ i.e. $Y = \sum_k^K \alpha_k u_k v_k'$ (a.1.2)

where U is an $n \times K$ matrix, D is an $K \times K$ diagonal matrix with elements $\alpha_1 \geq \alpha_2 \geq .... \geq \alpha_n \geq 0$, and V is a $K \times p$ matrix. The K columns of orthonormal n-vectors $u_1, u_1, ..., u_K$ of U, called the *left singular vectors*, are an orthonormal basis for the columns of Y and are the eigenvectors of $YY\prime$, with associated eigenvalues $\alpha_1^2, ..., \alpha_K^2$. Similarly the K columns of orthonormal p-vectors of $v_1, v_1, ..., v_K$, called the *right singular vectors*, are orthonormal basis for the (transposed) rows of Y and are the eigenvectors of $YY\prime$, with the same associated eigenvalues $\alpha_1^2, ..., \alpha_K^2$. The element $\alpha_1, \alpha_2, ..., \alpha_K$ of the diagonal matrix D_α are called the *singular values* of Y.

The matrix $\mathbf{F} \equiv UD_\alpha$ contains the *coordinates of the rows* of Y with respect to the basis vectors in V and the matrix $\mathbf{G} \equiv VD_\alpha$ contains the *coordinates of the columns* of Y with respect to the basis vectors in U.

In fact the matrix Y can be written as $\mathbf{Y} = FV'$, then the $i-th$ row y_i of Y can be written as: $y_i = \sum_k^K f_{ik} v_k$ (a.1.3). So that the $i-th$ row of F contains the coordinates of y_i with respect to the basis vectors in V . If the last terms of (a.1.2) corresponding to the smallest singular values are dropped then a least squares approximation of the matrix Y results. That is, if we define the matrix $Y_{[K^*]}$ as the first K^* terms of (a.1.2): $Y_{[k^*]} = \sum_k^{K^*} \alpha_k u_k v_k'$ (a.1.4) then $Y_{[k^*]}$ minimizes: $\|Y - A\|^2 \equiv \sum_i^n \sum_j^p (y_{ij} - a_{ij})^2$ (a.1.5) amongst all $n \times p$ matrices A of rank at most K^*. $\mathbf{Y}_{[K^*]}$ is called the rank K^* (least squares) approximation of Y.

The rows and the column of $Y_{[K^*]}$ are vectors contained in subspaces of dimension K^* which best fit the rows and the columns of Y in the sense of minimum sum of squared Euclidean distances. This solves the problem of minimizing the loss (a.1.1), that is the *ordinary principal components analysis*, when $w_1 = ... = w_n = 1$ and $D_p = I_p$.

Now let us consider the problem of minimization in presence of mass and dimension weight, this define the *generalized principal components analysis*.

[6] This type of distance function is often referred to as a diagonal metric.

Let us consider a $n \times p$ matrix $\tilde{\mathbf{Y}}$ of the deviation of $\mathbf{Y}$ from the mean vector $\bar{\mathbf{y}}$: $\tilde{\mathbf{Y}} \equiv \mathbf{Y} - \mathbf{1}\bar{\mathbf{y}}$. It can be decomposed as: $\tilde{\mathbf{Y}} = \mathbf{N}\mathbf{D}_\mu\mathbf{M}'$ (a.1.6) where $\mathbf{N}'\mathbf{\Omega}\mathbf{N} = \mathbf{M}'\mathbf{\Phi}\mathbf{M} = \mathbf{I}$, $\mathbf{\Omega}$ and $\mathbf{\Phi}$ being prescribed positive definite symmetric matrices. We denote this the *generalized singular value decomposition* in the metrics $\mathbf{\Omega}$ and $\mathbf{\Phi}$.

Let us now set $\Omega{=}D_w$ the masses and $\Phi{=}D_s$ the dimension weights.

Then the matrix : $\tilde{\mathbf{Y}}_{[K^*]} = N_{[K^*]}D_{\mu[K^*]}M'_{[K^*]} = \sum_k^{K^*} \mu_k n_k m'_k$, (a.1.7) minimizes: $\left\|\tilde{\mathbf{Y}} - \mathbf{A}\right\|^2 \equiv \sum_i^n \sum_j^p w_i s_j \left(\tilde{y}_{ij} - a_{ij}\right)^2 = \sum_i^n w_i \left(\tilde{y}_i - a_i\right)' D_s \left(\tilde{y}_i - a_i\right)$ (a.1.8) amongst all $n \times p$ matrices A of rank at most K^*, where $\tilde{y}_i$.and a_i are the rows of $\tilde{Y}$ and A respectively. From the form of the optimum, $\tilde{\mathbf{Y}}_{[K^*]}$, the columns vectors $m_1, m_2, ..., m_{K^*}$ of $M_{[K^*]}$ define the Φ-orthonormal basis for the optimal subspace Q^*. The coordinates of the vectors $\mathbf{y}_i - \bar{\mathbf{y}}$ with respect to this base are the rows of $\mathbf{F}_{[K^*]} \equiv N_{[K^*]}D_{\mu[K^*]}$(a.1.9).

The generalized SVD provides the required solution for any prescribed dimensionality K^*. Let us suppose to represent the vectors $\mathbf{y}_1, ..., \mathbf{y}_n$ in a subspace at one dimension $K^* = 1$, geometrically along a straight line. It means that the first singular vectors and the associated singular values (the largest) provide the optimal solution. If we want to represent the vectors $\mathbf{y}_1, ..., \mathbf{y}_n$ over a plane ($K^* = 2$), it means that the first and the second pairs singular vectors and associated singular values provide the optimal solution. In general, we could obtain K basis vectors, defined as *principal axes* of the rows of $\mathbf{Y}$.

The squared singular values give an idea of how well the matrix is represented along the principal axes. The total variation of a matrix $\tilde{\mathbf{Y}}$ is quantified by its squared norm, $\left\|\tilde{\mathbf{Y}}\right\|^2_{D_s D_w} \equiv \sum_i^n w_i \tilde{y}'_i D_s \tilde{y}_i = \sum_k^K \mu_k^2$ (a.1.8). Similarly the total variation of the approximation $\tilde{\mathbf{Y}}_{[K^*]}$ is: $\left\|\tilde{\mathbf{Y}}_{[K^*]}\right\|^2_{D_s D_w} = \sum_k^{K^*} \mu_k^2$ and the unexplained variation is $\left\|\tilde{\mathbf{Y}} - \tilde{\mathbf{Y}}_{[K^*]}\right\|^2_{D_s D_w} = \sum_{k=K^*+1}^{K^*} \mu_k^2$ which is minimized.

Since, $\tilde{y}_i = y_i - \bar{y}$ (a.1.2) is the weighted sum of squared distances of the vectors y_i from their centroid $\bar{y}$, a type of generalized variance which is defined *inertia* of the set of the vectors , or *total inertia*. Because the total inertia is the sum of the squared singular values of $\mathbf{Y}$, the k^{th} principal axis accounts for an amount μ_k^2 of the total inertia. Hence, the total inertia is decomposed along the principal axes.

In order to compute the generalized SVD of the matrix a general procedure is to pursue the following steps. Let us consider $\mathbf{B} = \mathbf{D}_w^{1/2}\tilde{\mathbf{Y}}\mathbf{D}_s^{1/2}$. Then find the ordinary SVD of $\mathbf{B}$: $\mathbf{B} = \mathbf{U}\mathbf{D}_\mu\mathbf{V}'$. Let $\mathbf{N} = \mathbf{D}_w^{-1/2}\mathbf{U}$, $\mathbf{M} = \mathbf{D}_s^{-1/2}\mathbf{V}$, $\mathbf{D}_\mu = \mathbf{D}_\alpha$. Hence, $\tilde{\mathbf{Y}} = \mathbf{N}\mathbf{D}_\mu\mathbf{M}'$ is the generalized SVD (Greenacre, 1984).

7.2 CANONICAL CORRELATION ANALYSIS

7.2.1 INTRODUCTION

In this section we shall describe the geometry of canonical correlation analysis[7] (henceforth, CCA) as equivalent approach to correspondence analysis[8]. It can be shown that to carry out the correspondence analysis of a contingency table is equivalent to find the canonical correlation of the disjunctive matrices associated to (Saporta, 1990).

Then we shall consider the generalisation of the method proposed by Carrol (1968) in order to show that the multiple correspondence analysis is a particular case.

Let $Z \equiv [Z_1; Z_2]$ be a matrix of dimension $I \times J$ which is divided into two blocks Z_1 and Z_2 of J_1 and J_2 variables respectively, such that $J = J_1 + J_2$.

The object of CCA is to capture the correlations between two sets of variables. It tries to find basis vectors for two sets of variables such that the linear correlations between the projections onto these basis vectors are mutually maximized.

Formally it can be expressed as finding linear combinations $u = Z_1 a$ and $u = Z_2 b$ of each set of variables which have maximum correlation ϱ[9]. The maximum number of correlations that it can be found is equal to the minimum of the columns dimension J_1 and J_2. A complete analysis consists of identifying further linear combinations $u_k = Z_1 a_k$ and $v_k = Z_2 b_k$ each uncorrelated with previous linear combinations $u_1 ... u_{k-1}$ and $v_1 ... v_{k-1}$, which have maximum correlation. If $J_1 \leqslant J_2$, it can be identified at most $k = J_1$ canonical correlations ϱ_k, $k = 1, ...J_1$ in descending order. The vectors $u_1 ... u_k$ and $v_1 ... v_k$ and are called *canonical score vectors* to which corresponds $a_1, ...a_k$ and $b_1, ...b_k$ *canonical weights*. The correlation between the *canonical score vectors* $\mathbf{u}$ and $\mathbf{v}$ is given by $\varrho = (a' S_{12} b) / [(a' S_{11} a)(b' S_{22} b)]^{1/2}$ (a2.1.1) where S_{12} is the covariance matrices between Z_1 and Z_2, S_{11} and S_{22} are the covariance matrices of Z_1 and of Z_1 respectively.

The first problem is finding the vectors of *canonical weights* that maximize equation (a2.1.1) above. It can be shown that the vectors a_k and b_k of *canonical weights* can be obtained from the left and right singular vectors of the matrix $S_{11}^{-1/2} S_{12} S_{22}^{-1/2}$ Specifically, if the SVD of the matrix is $S_{11}^{-1/2} S_{12} S_{22}^{-1/2} = W D_\varrho X'$ with $W'W = X'X = I$ (a2.2.2) then the singular values of the diagonal of D_ϱ are the canonical correlations, while the matrices $A \equiv [a_1 ....a_K]$ and $B \equiv [b_1 ....b_K]$ of *canonical weights* are simply $A = S_{11}^{-1/2} W$ and $B = S_{22}^{-1/2} X$.

The orthonormality of the singular vectors of W and X stated in (2.2.2)

[7] Canonical corelation analysis was introduced originally by Hotelling (1936).

[8] The geometry of canonical correlation is given by Greenacre (1984, section 4.4).

[9] Note that the correlation coefficient ρ only provides a measure of the linear association between the two sets of variables: when the two sets of variables are uncorrelated, i.e., when their correlation coefficient is zero, this only means that no linear function describes their relationship. A quadratic relationship or some other non-linear relationship are not ruled out.

implies that A and B are standardized as follows: $AS_{11}A = BS_{22}B = I$ (a2.2.3). This is a standardization that allows the vectors of *canonical scores* to be of unit variance (and uncorrelated).Thus (2.2.3) is a set of identification conditions on the scale of the canonical weights and equivalently of the canonical scores in each u_k and v_k.

In order to identify the origins of the vectors of *canonical scores*, their means are conventionally set to zero, which is equivalent to each variable (i.e. column) of Z_1 and Z_2 being centred with respect to its mean.

In case of data coded as a bivariate indicator matrix the application of CCA to such data breaks down because the covariance matrices S_{11}, S_{22} are singulars which implies that they cannot be inverted.

In fact if we let r and c be the vectors of the means of the columns of Z_1 and Z_2, respectively, D_r and D_c and be diagonal matrices of r and c respectively, then the covariance matrices are simply: $S_{11} = D_r - rr'$, $S_{22} = D_c - cc'$ and are of ranks $(J_1 - 1)$ and $(J_2 - 1)$ respectively. We can use one of a number of generalized inverses to carry the classical theory through. In this particular situation it turns out that the complete solution to the problem is contained in the CCA of the data without prior centring of the columns of Z_1 and Z_2. Variances and covariances are defined with respect to the origin (not the mean) and the analysis yields a trivial maximum solution where the canonical correlation is one and associated canonical weight vectors are ones (J_1 vectors of ones and J_1 vectors of ones), after which the canonical solutions are those of the centred problem. Here the non trivial solutions are centred and thus identified with respect to the origin by virtue of their uncorrelation with the trivial solution.

We can describe the geometry of CC in two different equivalent ways. The first way of describing CC is to think of the columns of Z_1 and Z_2 as points in I-dimensional Euclidean space.

In the case of quantitative data the columns of Z_1 and Z_2 are centred, which means that they have been projected orthogonally onto the $(I - 1)$ dimensional subspace orthogonal to 1. The set of all linear combinations of the columns of Z_1 and of Z_2 form J_1 and J_2 dimensional subspaces respectively, and the cosine of the angle between any two vectors is equal to their correlation. Hence CCA is the search for any two vectors, u and v, in these respective subspaces, which subtend the smallest angle. The procedure is repeated in the $(J_1 - 1)$ and $(J_2 - 1)$– dimensional subspaces orthogonal to the canonical score vectors u and v to obtain a second canonical correlation and score vectors. Clearly, if $J_1 \leqslant J_2$ then we would eventually end up with a set of J_1 canonical score vectors in each subspace, the first set explaining all J_1 dimensions of the column space Z_1, while the second set leaves $J_2 - J_1$ dimensions of the column space of Z_2 unexplained. In this framework the correlations between the new canonical variables and the original variables are simply the angle cosines between the canonical score vectors and the columns of Z_1 and Z_2.

When Z_1 and Z_2 are indicator matrices, the vector 1, being the sum of the columns of Z_1, and likewise of Z_2, is common to both subspaces. If we centre the columns, that is we project onto the subspace orthogonal to 1, would reduce the dimensionality of each subspace and we would not be able to identify the score vectors in this lost dimension. If we omit centring the columns, the highest canonical correlation is 1, when u and v are both collinear with the vector 1 and thus subtend a zero angle. This is the trivial solution. Subsequent canonical score vectors are orthogonal to 1 in each subspace, and are thus centred as required. If $J_1 \leqslant J_2$, there will be only $(J_1 - 1)$ non-trivial canonical correlation for such data.

The second way of studying canonical correlation analysis geometrically is to think of the rows of Z_1 and Z_2 as $z_1^1, z_2^1, ...z_I^1$ and $z_1^2, z_2^2, ...z_I^2$ vectors in the corresponding spaces of dimensionalities J_1 and J_2 respectively.

Let us define these spaces as Mahalanobis spaces since the metrics imposed on these ones are defined by the inverses of covariance matrices S_{11}^{-1} and S_{22}^{-1} conventionally called Mahalanobis metrics. The columns $S_{11}a_k$ of $S_{11}A$ and $S_{22}b_k$ of $S_{22}B$ of can be considered orthonormal basis vectors in these two spaces, since: $(S_{11}A)' S_{11}^{-1} (S_{11}A) = (S_{11}B)' S_{11}^{-1} (S_{11}B) = I$ which is equivalent to standardization (a2.2.3). The coordinates of the rows of Z_1 and of Z_2 with respect to the *canonical axes*, $S_{11}a_k$ and $S_{22}b_k$ respectively, are the corresponding vectors of the *canonical scores*: $Z_1 S_{11}^{-1} (S_{11}a_k) = Z_1 a_k = u_k$ and $Z_2 S_{22}^{-1} (S_{22}b_k) = Z_2 b_k = v_k$.

In the J_1 and J_2 - dimensional geometries of the rows of Z_1 and Z_2 respectively the variables are represented as axes with the canonical weights as coordinates.

We can combine these two viewpoints. From (a2.2.2) and (a2.2.3) we see that the two sets of coordinates $A_{(k^*)}$ and $B_{(k^*)}$ are related as follows: $S_{11}^{-1} S_{12} B_{(k^*)} = A_{(k^*)} D_{\varrho(k^*)}$(a2.2.4) $S_{22}^{-1} S_{12} A_{(k^*)} = B_{(k^*)} D_{\varrho(k^*)}$(a2.2.5) where $D_{\varrho(k^*)}$ denotes the $k^* \times k^*$ diagonal matrix of the first k^* canonical correlations $\varrho_1, ..., \varrho_k$.

7.2.2 GENERALISED CANONICAL CORRELATION ANALYSIS

The generalised canonical correlation analysis (GCCA) describe the correlations between K data matrices observed for the same n units. The method is a generalisation of CCA[10] and it requires a criterion for measuring the correlation between K data matrices. The criterion consists in finding *canonical variates* y_j so that the sum of the squared correlation coefficients between them and each data matrix is maximum. Let $Z = [Z_1; Z_2; ..; Z_q; ..., Z_Q]$ be a $n \times p$ matrix which consists of Q blocks of centred matrices Z_q[11], the objective that correspond to the j^{th} canonical variate can be written as: $s_j = \sum_{q=1}^{Q} y_j' Z_q \left(Z_q' Z_q\right)^{-1} Z_q' y_j$.The variates y_j are determined in order to maximize s_j with the constraints of being normalised and orthogonal to each other: $y_j' y_j = 1$ and $y_i' y_j = 0$ for $i \neq j$.

The criteria y_j are maximum when the *canonical variates* are eigenvectors of $n \times n$ the matrix S: $S = \sum_{q=1}^{Q} Z_q \left(Z_q' Z_q\right)^{-1} Z_q'$ (a3.1).

[10] The generalization of the method was proposed by Carrol (1968).
[11] Moreover the columns of Z_q should be linearly indipendents.

7.2.3 MULTICORRESPONDENCE ANALYSIS

Let us suppose that $Z = [Z_1; Z_2; ..; Z_q; ..., Z_Q]$ is a $n \times p$ matrix composed by Q indicators matrices Z_q. Let us define D_q a diagonal matrix of order $n \times p_q$ corresponding to the sum of the columns of the matrix Z_q. Let us define $D \equiv (D_q)$ a $p \times p$ diagonal matrix which is composed by Q diagonal blocks.

In this case the matrix (a3.1) is equivalent to: $S = \sum_{q=1}^{Q} Z_q (D_q)^{-1} Z'_q$ (a3.1.1) and the maximization problem is $s_j = y'_j S y_j$ with the constraint $y'_j y_j = 1$. Then the maximization conditions are $Sy_j = \lambda y_j$ and $\sum_{q=1}^{Q} Z_q (D_q)^{-1} Z'_q y_j = \lambda y_j$ (a3.1.2). Since for each $h < Q$ can be written that $Z'_h y_j = u_h$ and in general that: $\mathbf{Zy} = \mathbf{u}$ where $\mathbf{u}$ is a vector with p components so that u'=$\left(u'_1, u'_2, ..., u'_h, ..., u'_Q\right)$.

Multiply both members of (a3.1.2) for $Z'_h y_j$ then can be written that for $h = 1, ...Q$: $\sum_{q=1}^{Q} Z'_h Z_q (D_q)^{-1} Z'_q u_q = \lambda u_h$ (a3.1.3). The compact form of the (a3.1.3) is: $\mathbf{Z'ZD^{-1}u} = \lambda \mathbf{u}$

The MCA is $(1/q)\,\mathbf{Z'ZD^{-1}u} = \lambda \mathbf{u}$ where $\lambda = \mathbf{q}\lambda$ The eigenvalues found by GCCA is q times those found in CCA on the matrix $\mathbf{Z}$.