THE CROSS-SECTIONAL CHARACTERIZATION OF ACCOUNTING DATA

Duarte Trigueiros

ISCTE,
Av. Forças Armadas, 1600 Lisbon, Portugal
Phone: +351 (1) 793 50 00, Fax: +351 (1) 796 47 10,
E-Mail: dmt@sara.inesc.pt

April 20, 1992

Introduction

This study develops a statistical basis for financial statement analysis. It introduces a unified view able to clarify issues widely discussed in the accounting literature: The validity of ratios, the observed variety of their statistical distributions, how to overcome their known drawbacks when used as input variables in statistical models. It is based on the observation of the cross-sectional characteristics of a large set of financial statements from firms in the UK.

The validity of ratios in financial statement analysis: Accounting reports are an important source of information for managers, investors and financial analysts. Ratios are the usual instruments for extracting this information. They are supposed to control for the effect of size and to highlight some noteworthy features of the firm. However, the widespread use of ratios in financial statement analysis has been frequently criticised in the accounting literature. The main sources of concern are the possible existence of functional relations between the numerator and the denominator of ratios which may be non-proportional or non-linear. In both cases, ratios would introduce a bias when controlling for size. Whittington (1980) [30] discusses this subject. A recent review of the published research and discussions can be found in Berry and Nix (1991) [4].

Ratios as input variables for statistical modelling: Statistical techniques are also used to extract information from databases gathering accounting reports and related outcomes. The goal is to construct models suitable for prediction or for isolating the main features of the firm. An early model is that of Beaver (1966) [3] who used accounting ratios to predict financial distress. Many researchers followed him, using more sophisticated tools. Linear Regressions and Fisher's Multiple Discriminant Analysis are the most popular algorithms. Logistic Regression can also be found in some studies. Foster (1986) [15] offers a review of the modelling practice involving financial statements.

The use of ratios in statistical models seems to be an extrapolation of their use in financial statement analysis. However, there are difficulties involved in using ratios as statistical variables. Firstly, the distribution of ratios is of many different sorts: Most ratios exhibit strong positive skewness, but some are Gaussian or even negatively skewed. Hitherto, no explanation has been given to this variety. Secondly, ratios exhibit heteroscedasticity and cases which are severely influential. Finally, it is difficult to decide which ratio should be selected for a given task. As N accounting items can generate up to $N^2 - N$ ratios, some research seems to get lost in a prolific use of all sorts of combinations of items. Chen and Shimerda [8] (1981) review this topic.

In order to overcome the above problems, accounting modelling practice rely on general-purpose recipes: Improvements in normality are sought by pruning out tails and empirically trying different transformations. It's common practice to mix up in the same model square root and log transformations.

Heteroscedasticity is treated as a separate phenomenon and requires further manipulation, typically a weighting of cases. Finally, the multicolinearity generated by the excess of variables is also viewed as an accident in its own right so that the measures recommended for general cases are applied — A few Principal Components are extracted and used instead of the original variables. The model parameters, after this pruning, scaling and rotating, are difficult to interpret.

The basis of this study: We argue that any attempt to build a basis for financial statement analysis and statistical modelling by looking into financial ratios is not likely to succeed. Ratios are bivariate relations. Their behaviour is determined by that of their components, plus the interaction between them. In this study we begin by the opposite end. We show that the nature of accounting information becomes greatly clarified after explaining the cross-sectional behaviour of raw accounting numbers. And when doing so, also the problems quoted above — non-proportionality, transformations versus the prunning of outliers, heteroscedasticity, multicolinearity — are solved.

Contents: Chapter 1 introduces the data. Then, it empirically assesses the statistical nature of raw accounting numbers. Both individual industries and overall samples are examined. It is shown that raw data, unlike ratios, have a regular and predictable distribution.

Chapter 2 studies the regularities devised in chapter 1 and extracts the most immediate consequences for the modelling of accounting relations. The existence of outliers and heteroscedasticity in statistical models is explained. The use of regressions instead of ratios is discussed.

Chapter 3 examines the joint behaviour of more than one item. It is shown that items belonging to the same report share most of their variability. This chapter also discusses the statistical nature of items which are a subtraction of two items.

Chapter 4 examines the main sources of concern regarding the validity of ratios. New models, which are extensions of ratios able to account for non-linearity and non-proportionality, are introduced.

Chapter 5 studies the distribution of ratios. The main rules governing their cross-sectional behaviour are described. Consequences for the statistical modelling of accounting data are extracted.

Chapter 6 is about size and industrial grouping as the main sources of variability present in the data. A method for isolating size as a statistical effect is developed and discussed. The homogeneity and complexity of some industrial groups are also assessed.

Finally, chapter 7 studies the content on information of ratios. Given the two components of a ratio, it is shown that all the size-adjusted information conveyed by them can be expressed in terms of the ratio itself plus a remainder. This chapter shows that such a remainder contains, in some cases, valuable information for financial statement analysis.

Acknowledgements: The experiments described in this study and most of the developments that followed, were carried out in the University of East Anglia in the UK. We thank Professor Robert H. Berry for his valuable support. We are also grateful to Professor Geoffrey Whittington and to Archie Pitts for their reading and suggestions. Professor S. Sudarsanam offered useful suggestions regarding a better understanding of the contents.

Contents

Introduction iii

1 The Distribution of Raw Accounting Numbers 1

1.1 The Method 2
1.2 Items and Industrial Groups Examined 4
1.3 Results 5
1.3.1 All Groups Together 5
1.3.2 One Industry at a Time 6
1.4 Other Possible Parameterization 7
1.4.1 Root Transformations 7
1.4.2 The Pareto Hypothesis 7
1.5 Assessing the Importance of Multiple Tests 9
1.6 Discussion and Conclusions 11

2 Outliers, Heteroscedasticity and Trimming 19

2.1 Outliers 20
2.2 Heteroscedasticity 20
2.3 Trimming 22
2.4 Conclusion 25

3 The Overall Effect of Size 27

3.1 Co-Variance Matrices in Log Space 27
3.2 The Subtraction of Two Items 30
3.2.1 How to Model With Samples Having Negative Cases 31
3.2.2 Alternative Variables 35
3.3 Summary 35

4 Validity and Extension of Ratios 36

4.1 Ratios and the Common Effect 37
4.2 Free-Slope Ratios for Prediction 39
4.3 Non-Proportional Ratios 41
4.4 The Basis for the Existence of Non-Proportionality 44
4.5 Using Extended Ratios 45
4.6 Discussion and Conclusions 47

5 The Distribution of Ratios **49**

5.1 The Effect of External Constraints 49

5.2 Comparing Constrained and Non-Constrained Ratios 51

5.3 The Log-Leptokurtosis of Ratios 53

5.4 The Two Sources of Statistical Behaviour 57

5.5 Avoiding Asymmetry in Ratios 57

5.6 Summary 59

6 Size And Grouping **60**

6.1 Estimating Size: The Case-Average Model 60

6.2 The Homogeneity of Industrial Groups 64

6.2.1 Introduction and Related Research 64

6.2.2 Measuring the Significance of a Grouping 65

6.2.3 Assessing the Complexity of Grouping Effects 67

6.3 Summary 71

7 The Complement of a Ratio **72**

7.1 Assessing the Remainder 72

7.2 Is the Remaining Information Useful? 73

7.3 Creating Bivariate Tools 74

7.3.1 Non-Rotated Plots 74

7.3.2 Rotated Plots 76

7.3.3 Financial Diagnostics and the Rotated Residual Plot 78

7.4 Using Complementary Ratios 80

7.4.1 How to Build the RRP 80

7.4.2 Reading the RRP 81

7.5 Summary 86

8 Conclusions and Tables **87**

Chapter 1

The Distribution of Raw Accounting Numbers

Accounting numbers, as found in databases containing collections of annual reports of firms, can be viewed as statistical variables. Each firm is a case. For a given item, say, Fixed Assets or Sales, a particular collection of firms forms a cross-sectional sample.

In this chapter we show that the probability density function governing raw accounting numbers is much more predictable and regular than the one of ratios. The lognormal distribution emerges as a general characteristic of the observed items. We also explore several other distributions and we show that they are not plausible. In order to assess the importance of the results of the performed tests we compare them with simulated ones. We also examine individually the few cases of departures from lognormality.

We use data from British firms belonging to 14 industrial groups in four broad areas: Engineering, Processing, Textiles and Food Manufacturers. Both individual industries and overall samples are examined. Our study contemplates a period of five years (1983-1987) but it is a cross-sectional study: Each year is studied individually.

Previous research: Since practitioners use ratios, not the raw data, the statistical properties of items received little attention in the literature. Ratios, of course, have been the object of a much bigger effort. Horrigan [18] (1965) is an early work on this subject. He analyzed 17 ratios for 50 companies over the period 1948-57 reporting positive skewness. Horrigan explained it as a result of effective lower limits of zero. O'Connor [25] (1973) discovered that for all his 10 ratios in a set of 127 companies during the period 1950-66, skewness prevailed once again. Also Bird and McHugh [5] (1976) analyzed 5 ratios for 118 firms over the period 1967-71 in Australia finding positive skewness. But they considered it as an accident and implicitly suggested the pruning or winsorizing of distributions until achieving normality.

The Deakin study [10] (1976) is frequently quoted. It showed that the positive skewness could not be ignored in his sample of 11 ratios for the period 1955-73. He concluded that

> ..."as a result of this analysis it would appear that assumptions of normality for financial ratios would not be tenable except in the case of the Total Debt to Total Assets ratio."

The Bougen and Drury study [6] (1980) was based on UK. firms. It collected data on 700 industrial firms (1975) and analyzed 7 ratios, concluding that skewness could not be ignored. Also Buijink [7] (1984) reported the persistency of skewness over a large period. Barnes [2] (1982) argued that skewness on ratios could be the result of deviations from proportionality between the numerator and the denominator. Frecka and Hopwood [16] (1983) extended Deakin's 1976 study for a longer period and reported similar findings. They also tried to achieve normality by applying square root transformations and pruning the remaining outliers, proposing such procedure as a way of dealing with the problem of deviations from normality. Ezzamel and Mar-Molinero wrote a recent report (1990) on the distribution of ratios using UK data [13]. These authors also investigate the effects of a family of transformations in the distribution of ratios.

Two studies by McLeay [23] [24] (1986) refer to items, not to ratios. McLeay distinguishes two broad classes of items: Sums of similar transactions, whose sign remains the same; and differences, which could be of either sign or zero. McLeay argues that the former ones ought to be lognormal since they are related to the size of the firm which could be a stochastic process governed by the Gibrat law (see section 1.4.2 and 4.1). McLeay accepts a qualitative difference between these two families of items.

1.1 The Method

In order to test the lognormality of accounting items we use a logarithmic transformation. We then apply to the transformed data the Shapiro-Wilk test of normality in an improved version due to Royston [26] (1982). This test can cope with large or small sample sizes and is recommended as a superior omnibus test. It has been used by Ezzamel, Mar-Molinero and Beecher [13] (1987) to test the normality of ratios. Berry and Nix [4] (1991, p. 110) discuss it in more detail. The Shapiro-Wilk test yields a statistic, W, ranging from zero to one. Values of W approaching 1 mean increasing normality. The significance, P, of W is the probability of obtaining such a W when the population is Gaussian.

The transformation we apply to an item called x isn't just the log of x. The lognormal distribution can have three parameters: The log mean and standard deviation, corresponding to the Gaussian ones, plus a third parameter accounting for overall displacements of the distribution. When a displacement of x, say $x - \delta$, and not x itself, is Gaussian after a log transformation, the distribution of x is known as Three-Parameter Lognormal. The range of x is thus $\delta < x < \infty$. The usual, two-parametric, lognormal distribution can be seen as a special case for which $\delta = 0$. Since δ is a lower bound for x, it is known as the threshold of the distribution. An introduction to the lognormal distribution can be found in Aitchison and Brown [1] (1957).

Lognormal distributions often are three-parametric. Thresholds play an important role in the understanding of non-proportionality in ratios, as we shall see in section 4.3.

Estimating the threshold: Some procedures available to estimate δ are also described by Aitchison and Brown. In our case, δ is estimated using a modified version of the procedure suggested by Royston [26] (pp. 123). His method consists of discovering by trial which δ maximizes the Shapiro-Wilk's W. In our case, the threshold is estimated as the smallest δ able to attain a significant W.

Figure 1 shows two examples. The significance, P, of W, improves when we subtract a constant small δ to all observations in a three-parametric distribution, before transforming the data. By trying increasing δ we find an optimal W or $P(W)$. In the case displayed on the right, $\delta = 400$ is the value beyond which $P(W)$ no longer improves. Royston takes this $\delta = 400$ as the third parameter to introduce in the log transformation. On the left, we see how a small δ of -90 enhances the lognormality of Working Capital. We found sharp optimal values for W, making it possible to estimate δ in this way. However, as referred to above, the use of simulated thresholds showed us that the δ estimated by the Royston method are too large. To avoid over-estimation, thresholds must be estimated as the smallest δ able to render the sample two-parametric: $P(W) \geq 0.05$. Figure 1 also shows values for thresholds estimated in this way.

Notice that the use of δ cannot be considered as similar to the practice of adding a constant to all the cases in a sample to avoid negative-valued cases. Here, we don't allow δ to change the sign of cases.

1.2 Items and Industrial Groups Examined

We examined 18 items. They are listed in table 1 (page 14). There are 11 items from the Balance Sheet, 5 from the Profit and Loss Account, 1 from the Sources and Applications of Funds statement, and one which is not standard. These items are frequent as components of ratios. The selection of EX (Operating Expenses less Wages), is intended to get a picture of the cost structure of firms, using disclosed data. The inclusion of N (Number of Employees) concerns findings discussed later on: This variable conveys important information regarding the classification of industrial groups. It is also useful to compare accounting items with non-accounting variables exhibiting similar statistical behaviour.

In our study this set of selected items can be divided in two groups.

1. Those having only positive-valued cases, as Sales or Inventory.

2. Those items which can have both positive and negative-valued cases in the same sample, as Earnings and Working Capital. In this group we perform, when possible, two tests of lognormality:

   (a) Using only the positive-valued cases in each sample, and

   (b) using only the absolute value of the negative-valued cases. For small samples this test wasn't carried out since the number of cases was too small.

The reason for performing separated tests is that cross-sections of positive and negative-valued cases shouldn't be mixed up in the same sample as they represent different situations. Lev and Sunder [22] (1978, p. 201) explain in detail the reasons for avoiding such mixed samples.

It seems as if, by testing the sets of positive and negative-valued cases separately, we break the continuity of the sample and lose the information describing the passing through the zero value. In fact, we lose nothing by using these split samples in cross-sectional studies. Cross-sections are not about the time-history of one unique object. They are about many objects at the same instant in time. In section 3.2 we discuss in more detail the applicability of these split cross-sectional samples.

The groups: All companies quoted on the London Stock Exchange are classified into industrial groups according to the Stock Exchange Industrial Classification (SEIC). The tested samples were drawn from the Micro-EXSTAT database of company financial information provided by EXTEL Statistical Services Ltd, which covers the top 70% of UK industrial companies. We selected 14 manufacturing groups (see table 2 on page 14). This set of industries is the one examined by Sudarsanam and Taffler [29] (1985) in their study of the separability of industries, based on accounting information.

Two kinds of samples were examined:

All Groups Together, in which the 14 industrial groups are gathered in one unique sample.

One Industry at a Time, for samples of only one industrial group.

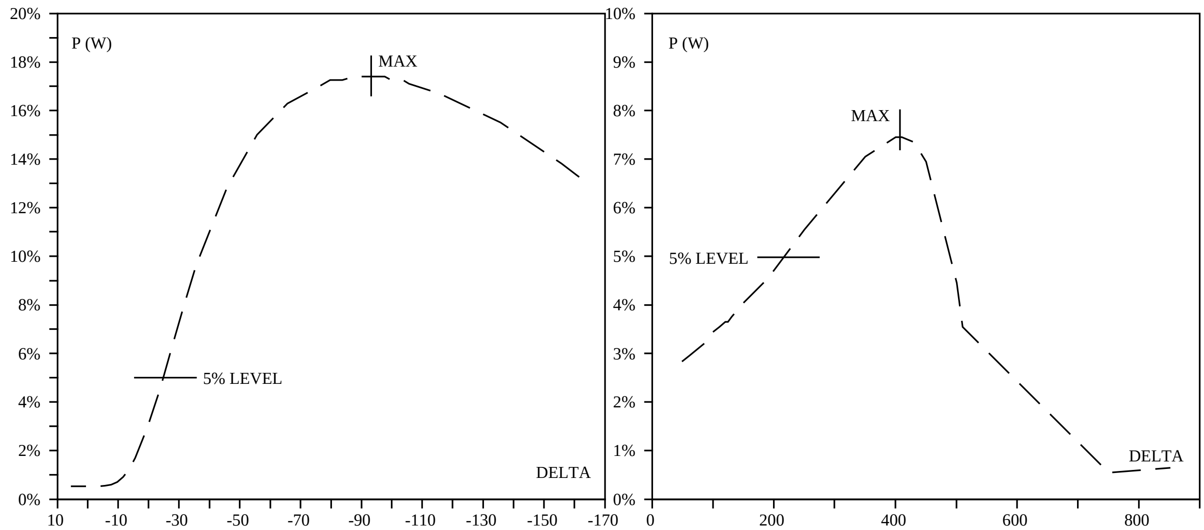

Figure 1: The significance, (P), of the Shapiro-Wilk's W for varying deltas. In these two cases we would accept a three-parametric lognormal distribution with $\delta \approx -20$ (left) and $\delta \approx 2000$ (right).

TA	Total Assets	NW	Net Worth
FA	Fixed Assets	DEBT	Long Term Debt
D	Debtors	C	Creditors
CA	Current Assets	CL	Current Liabilities
I	Inventory	TC	Total Capital Employed
WC	Working Capital		
EX	Operating Expenses less Wages	S	Sales
EBIT	Earnings Before Interest and Tax	W	Wages
OPP	Operating Profit		
FL	Gross Funds From Operations	N	Number of Employees

Table 1: List of accounting items examined by this study and their abbreviations.

PROCESSING:	14	Building Materials	32	Metallurgy
	54	Paper and Packing	68	Chemicals
ENGINEERING:	19	Electrical	22	Industrial Plants
	28	Machine Tools	35	Electronics
	41	Motor Components		
TEXTILES:	59	Clothing	61	Wool
	62	Miscellaneous Textiles	64	Leather
FOOD:	49	Food Manufacturers		

Table 2: List of the industrial groups examined by this study and their SEIC number.

Each test is performed five times for reports from 1983 to 1987. None of the companies present in the database was excluded.

1.3 Results

In this section we display the results of testing the lognormal hypothesis for two kinds of samples. Firstly, the large ones containing all the 14 industries together. Secondly, the small ones drawn from one industry at a time. These two groups of tests represent two possible levels of homogeneity worth exploring. One single group, if homogeneous, yields homogeneous samples. Two or three groups, each of them homogeneous, can yield samples which are severely non-homogeneous. But 14 groups in the same sample, all of them sharing a common attribute, are likely to apportion random effects rather than fixed ones, as the grouping acts as a random variable itself. In that case, a second level of homogeneity could be attained by gathering all industries in one sample. The examination of such overall samples is interesting as they represent the common attribute they share.

1.3.1 All Groups Together

The large samples mentioned above are now examined. The common attribute to consider here as a possible source of homogeneity is the industrial character of all the gathered firms.

Positive-valued cases: In tables 3 on page 15 and 4 on page 16, we display the number of cases in each sample and the statistics obtained when applying the Shapiro-Wilk test, along with some usual measures of normality (kurtosis and skewness) to the 13 positive-valued accounting items and to the positive values of the 4 items having both positive and negative cases. We also included Long Term Debt for which only the non-zero cases were selected. When nothing is said, items yielded non significant departures from two-parametric lognormality. When there is a significant departure, the significance is displayed. In all the significant departures observed, the introduction of a small threshold made it vanish.

The results show that 11 of the 18 items are two-parametric lognormal during the period 1983-1987. Sales and Operating Expenses less Wages, Net Worth, Debtors, Fixed Assets, Inventory and Total Capital Employed, along with the positive-valued cases of Earnings, Operating Profit, Long Term Debt and Working Capital, are persistently two-parametric. The remaining 7 variables are either two-parametric or three-parametric depending on the year. None is persistently three-parametric during five years.

The values that the threshold, δ, assumes whenever a three-parametric transformation is required, are near the smallest case in the sample. The skewness and the kurtosis of raw data is so high that its computation causes problems. After the log transformation, the skewness stabilizes. However, all the samples exhibit, after transformed, some residual leptokurtosis.

Negative-valued cases: We also checked the negative values of items having both positive and negative-valued cases. We selected the set of negative cases and then we applied logs to their absolute values. Table 5 (page 17) shows the results. Operating Profit and Working Capital are two-parametric for the whole period. Earnings and Gross Funds From Operations are, in one or two years, three-parametric.

Item	Year	1983	1984	1985	1986	1987
	N. Cases	555	649	677	702	688
Sales	SKEW	0.058	-0.035	0.126	0.107	-0.08
	KURT	0.765	0.842	0.604	0.39	0.854
	W	0.982	0.983	0.984	0.983	0.982
	$P(W)$					
	best W					
Net Worth	SKEW	0.174	0.163	0.253	0.289	0.289
	KURT	0.468	0.402	0.255	0.263	0.353
	W	0.987	0.985	0.984	0.983	0.983
	$P(W)$					
	best W					
Wages	SKEW	0.322	0.246	0.384	0.35	0.239
	KURT	0.177	0.252	0.067	0.091	0.266
	W	0.978	0.984	0.975	0.977	0.98
	$P(W)$	0.015		0.001	0.01	0.02
	best W	0.988		0.988	0.989	0.981
Inventory	SKEW	-0.089	-0.117	0.019	-0.202	-0.302
	KURT	0.69	0.577	0.572	1.217	1.331
	W	0.985	0.986	0.989	0.986	0.985
	$P(W)$					
	best W					
Debtors	SKEW	0.052	-0.003	0.066	0.126	-0.036
	KURT	0.309	0.411	0.386	0.318	0.76
	W	0.984	0.987	0.986	0.987	0.991
	$P(W)$					
	best W					
Creditors	SKEW	0.285	0.176	0.236	0.242	0.196
	KURT	0.307	0.331	0.356	0.289	0.369
	W	0.978	0.983	0.981	0.979	0.982
	$P(W)$	0.014		0.06	0.006	
	best W	0.985			0.985	
Fixed Assets	SKEW	0.097	0.177	0.119	0.124	0.159
	KURT	0.421	0.11	0.114	0.113	-0.008
	W	0.988	0.983	0.984	0.987	0.984
	$P(W)$					
	best W					
Total Assets	SKEW	0.301	0.351	0.404	0.343	0.425
	KURT	0.546	0.276	0.228	0.349	0.309
	W	0.983	0.979	0.978	0.979	0.978
	$P(W)$		0.01	0.005	0.01	0.005
	best W		0.985	0.987	0.983	0.984
Current Assets	SKEW	0.237	0.349	0.056	0.295	0.345
	KURT	0.372	0.374	1.84	0.345	0.48
	W	0.982	0.984	0.985	0.98	0.979
	$P(W)$				0.03	0.02
	best W				0.985	0.985

Table 3: Lognormality of all groups together. First table. “best W” shows the value of Shapiro-Wilk’s W after introducing an appropriate threshold.

(cont.)		1983	1984	1985	1986	1987
Current	SKEW	0.26	0.155	0.21	0.273	0.262
Liabilities	KURT	0.366	0.446	0.462	0.36	0.417
	W	0.979	0.983	0.984	0.981	0.982
	$P(W)$	0.026			0.04	
	best W	0.986			0.988	
Number	SKEW	0.171	0.191	0.283	0.221	0.159
of	KURT	0.282	0.181	0.187	0.251	0.346
Employees	W	0.981	0.982	0.98	0.981	0.981
	$P(W)$			0.02	0.04	0.05
	best W			0.985	0.985	0.983
Expenses	SKEW	0.093	-0.108	-0.043	0.012	-0.124
	KURT	0.327	0.738	0.742	0.29	0.641
	W	0.981	0.986	0.988	0.985	0.984
	$P(W)$					
	best W					
Total	SKEW	0.28	0.156	0.322	0.338	0.34
Capital	KURT	0.35	0.449	0.126	0.18	0.282
Employed	W	0.9828	0.9854	0.9819	0.9823	0.9815
	$P(W)$					
	best W					
EBIT	SKEW	-0.061	0.094	0.165	0.232	0.305
	KURT	0.678	0.244	0.409	0.402	0.451
	W	0.9846	0.9887	0.9841	0.9854	0.9823
	$P(W)$					
	N. Cases	514	606	629	645	641
Operating	SKEW	-0.097	-0.053	-0.13	0.128	0.176
Profit	KURT	0.68	0.36	0.81	0.215	0.469
	W	0.9898	0.9918	0.9854	0.9843	0.9848
	$P(W)$					
	N. Cases	497	589	615	627	619
Long Term	SKEW	-0.106	-0.049	0.029	-0.01	-0.059
Debt	KURT	0.095	-0.023	-0.101	-0.15	-0.016
	W	0.9868	0.9842	0.9839	0.9857	0.985
	$P(W)$					
	N. Cases	358	439	479	518	510
Gross Funds	SKEW	0.084	0.049	0.228	0.176	0.1
From	KURT	0.492	0.448	0.286	0.4	0.938
Operations	W	0.9867	0.9872	0.9842	0.9833	0.98
	$P(W)$					0.026
	N. Cases	527	625	647	666	650
Working	SKEW	-0.093	0.215	0.103	0.061	0.288
Capital	KURT	0.487	-0.062	0.532	0.269	0.311
	W	0.9926	0.9839	0.9881	0.9850	0.9807
	$P(W)$					0.052
	N. Cases	505	587	610	641	626

Table 4: Lognormality of all groups together. Second table. “best W” shows the value of Shapiro-Wilk’s W after introducing an appropriate threshold. When the number of cases is different from the one displayed in previous table, it is signaled here.

1.3.2 One Industry at a Time

The lognormality of our set of items cannot be rejected for samples drawn from one industry. The detailed results of applying the Shapiro-Wilk test to all samples — the eighteen observed items for fourteen industries during five years — can be found in appendix (tables 24 to 29, on pages 89 to 94). Here we present two condensed tables.

Positive-valued items by industry: For the selected 14 industries, 13 positive-valued items were checked for lognormality with the Shapiro-Wilk test. This procedure was repeated for five years (1983-1987). Therefore the number of different samples tested was 910. Lognormality was observed for the generality of cases. A total of 793 samples (87.1%) yielded two-parametric lognormality. Table 6 on page 17 shows, by industry and by item, the number of departures during the considered period. As an example, the value 2 in column "FA" and raw "Building Mat.", means that departures from two-parametric lognormality were observed twice in five years for Fixed Assets, in the Building Materials group.

Only in one case (Wages in the Electronics industry) we obtained a persistent $P < 0.05$ for all the five years. Electronics is also the group having more such cases (almost 40%). Next comes Food Manufacturers and Industrial Plants with more than 30%. Groups like Paper & Packing, Chemicals, Machine Tools, Building Materials and Clothing have less than 10% of departures. Miscellaneous Textiles, Wool, Motor Components, Leather and Metallurgy have no departures at all.

Positive values of items having both positive and negative cases by industrial group: In this paragraph we report on the behaviour of the positive values of four items having both positive and negative cases like Working Capital or Earnings. We also include one item, Long Term Debt, for which only the non-zero cases were selected. That makes a total of 350 samples to be tested: Five variables for the same fourteen industries during a period of five years.

A total of 311 samples (88.9%) yielded no significant departure from the two-parametric lognormal hypothesis. The results by industrial group and by item are similar to those of the previous paragraph (see table 7, page 17). Again, items had a more homogeneous behaviour than industrial groups.

Negative values of items having both positive and negative cases by industrial group: It is impossible to test the lognormality of all industries when considering absolute values of negative cases. The size of the resulting samples is too small. Only two groups were large enough to provide enough negative cases and they were two-parametric lognormal.

Departures from the two-parametric hypothesis: We measured the number of times a significant departure from a two-parametric lognormal distribution was observed in a given sample for the period of five years. In 69% of the tests performed in samples having only positive values there is no departure at all during the five-years period. One departure or two can be observed in 21% of the cases. Three departures only occur in 5% of the cases, and four in 3% of the cases. As mentioned before, only once (Wages of the Electronics industry) the departure from the two-parametric assumption occurs during the whole period of five years. In appendix (page 88), we display the results in detail.

Tables 6 and 7 show that the industrial grouping mostly determines whether a sample is two or three-parametric lognormal. The item being examined is less important than the industry in explaining significant thresholds.

Item	Statistic	1983	1984	1985	1986	1987
EBIT	W	0.9572	0.9838	0.988	0.951	0.946
	sig W				0.042	0.048
	N. Cases	40	42	48	57	47
Operating Profit	W	0.9537	0.9851	0.9925	0.9686	0.9596
	sig W					
	N. Cases	57	58	62	74	69
Gross Funds From Operations	W	0.9594	0.9649	0.9792	0.9066	0.9644
	sig W				0.005	
	N. Cases	27	24	30	36	38
Working Capital	W	0.9640	0.9743	0.9647	0.9678	0.9609
	sig W					
	N. Cases	50	61	67	61	62

Table 5: Results of applying the Shapiro-Wilk test to the absolute values of negative cases. All groups together. Departures from the two-parametric assumption have sig W < 0.05.

Industry / Items	S	CL	TA	N	TC	CA	EX	W	C	NW	D	I	FA	Total by Ind.
Leather														
Metallurgy														
Motor Compon.														
Misc. Textiles														
Wool														
Clothing	1						1							2
Building Mat.										1			2	3
Machine Tools				1		1					1	1		4
Chemicals		3							1			1	1	6
Paper & Pack.	1					1	3				2			7
Electrical	1		3	2	4	2		1		2		3		18
Food Manufac.	4	3	3		2	2	4	2	4	1				25
Ind. Plants	4	2	2	3	2	1	1	1	2	2	2	1	2	25
Electronics	3	3	3	4	2	2		5	1	2	2			27
Total by Item	14	11	11	10	10	9	9	9	8	8	7	6	5	117

Table 6: Number of samples yielding a significant departure from two-parametric lognormality during five years, by industry and by item. Items having only positive values.

Industry / Items	EBIT	OPP	WC	DEBT	FL	Total by Ind.
Machine Tools						
Misc. Textiles						
Metallurgy						
Leather						
Electrical					1	1
Paper & Pack.			1			1
Building Mat.	1				1	2
Wool			1	1		2
Motor Compon.				1	1	2
Clothing		1			2	3
Chemicals				2	1	3
Electronics	2	2	2		1	7
Ind. Plants	1	2	2	1	1	7
Food Manufac.	2	1	1	4	3	11
Total by Item	6	6	7	9	11	39

Table 7: Number of samples yielding a significant departure from two-parametric lognormality during five years, by industry and by item. Items having positive and negative values.

Item		Square Root	Cubic Root	Fourth Root	Sixth Root	Eighth Root	Log
SALES	Skewness	3.59	2.43	1.86	1.32	1.07	0.412
	Kurtosis	15.8	7.97	4.93	2.62	1.76	0.091
	Shapiro-Wilk's W	0.627	0.778	0.847	0.904	0.926	0.968
WAGES	Skewness	4	2.84	2.26	1.7	1.45	0.777
	Kurtosis	19.4	10.2	6.61	3.82	2.72	0.488
	Shapiro-Wilk's W	0.56	0.711	0.785	0.851	0.879	0.94

Table 8: Comparing root transformations with the log one. Electronics, 1985.

Departures from the lognormal hypothesis: Twenty tests yielded a very small significance of W, denoting severe departures from lognormality. These bad samples didn't improve with the three-parametric transformation. We examined each one of them (see figure 28 in appendix). Seven of these samples had extreme outliers, clearly erratic. The other 13 were non-homogeneous, showing two or three clusters of firms. The non-homogeneous samples occur in two industries only — Food Manufacturers and Electronics — mainly affecting Sales, Wages, and the Number of Employees.

Our results suggest that, for small samples, lognormality is conditional on the homogeneity of the industry. The underlying stochastic mechanism determining the distribution of raw data has a trend towards lognormality, even when a particular sample is non-homogeneous.

1.4 Other Possible Parameterization

The logarithmic transformation can be viewed as a way of controlling for the skewness in a distribution. It makes sense to ask if the reduction in skewness achieved with logs is the appropriate one. In case less reduction is required we should use a square root or another root. In case more reduction is required we should use the Pareto distribution or another of its class.

1.4.1 Root Transformations

We tested a scale of roots progressively approaching the effect of a log transformation. For each of them we measured the skewness, kurtosis and W. Table 8 shows the results for two particularly badly-behaved groups. The signs of normality increase with increasing roots, achieving much better values with logs. In figure 2 (page 8) we display the graphical evolution of a frequency distribution when roots of increasing exponents were used to transform the data.

Characteristic behaviour of raw data: When assessing the distribution of ratios it is usual to find cases in which no transformation seems to improve the normality of the data. However, in the case of raw data, the situation is different. Not only the log transformation emerges as the most appropriate one, no matter the item or the industry. There is also a progress towards normality for roots of increasing exponent. The signs of normality increase with increasing roots and are optimal with logs. Contrasting with this, it is frequent to find in the literature references to an unpredictable outcome in the distribution of ratios after applying transformations such those we use here (see, for example, Ezzamel, Mar-Molinero and Beecher [13], 1987, pp. 473 to 476).

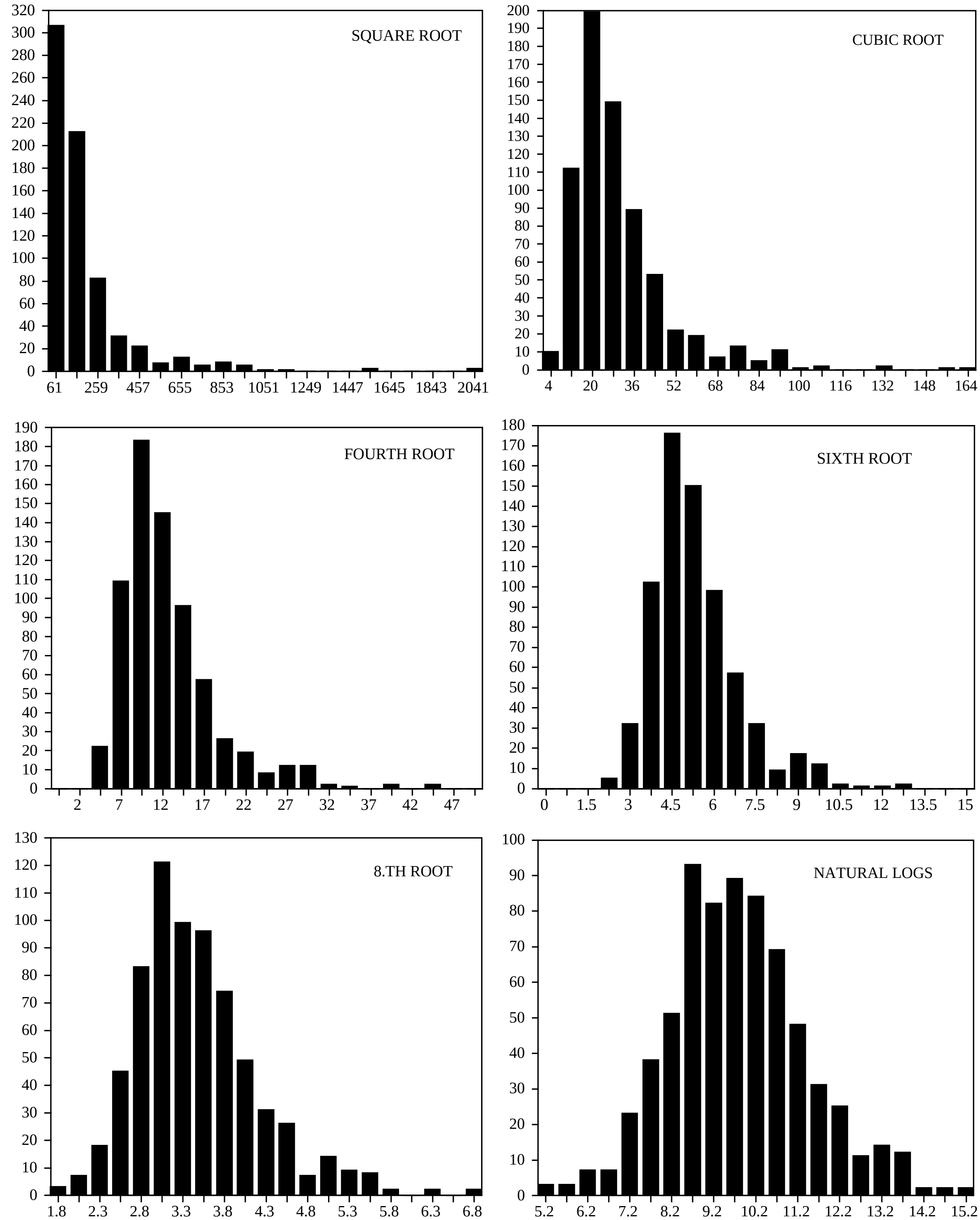

Figure 2: Evolution of the frequency distribution of Current Assets, 1986, all groups together, when several root transformations and a log one were applied.

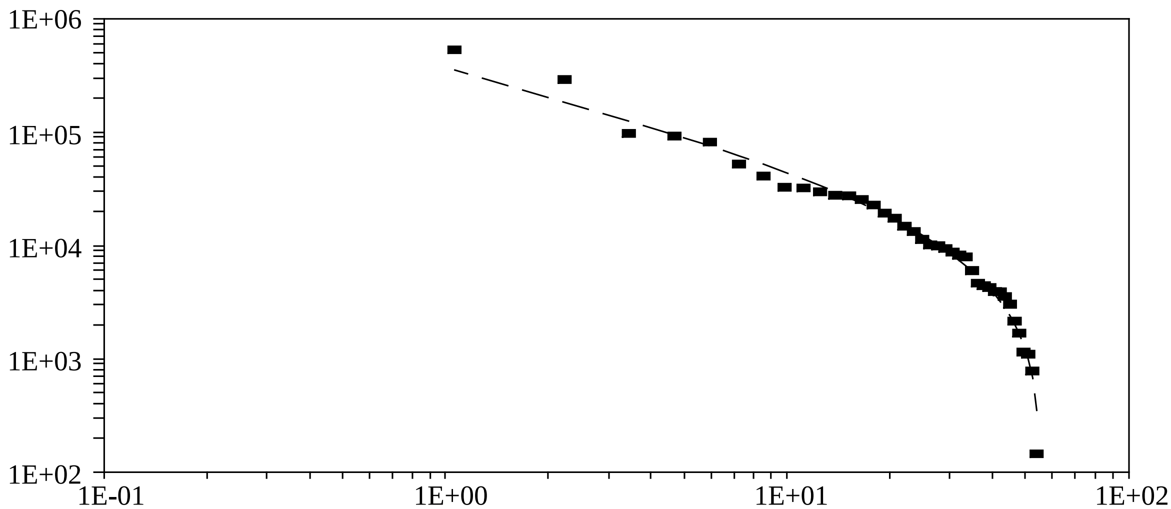

Figure 3: Pareto processes would exhibit a linear relationship between log ranks and log values. The dashed line is the corresponding lognormal deviate.

1.4.2 The Pareto Hypothesis

After discarding transformations less effective than logs in neutralizing skewness, we tried a more powerful one. Pareto processes have cumulative distributions for which the relation between the observed values and the rank is logarithmic. If x is a random variable governed by a Pareto process we observe

$$\log x = \log M - \beta \times \log r$$

in which r is the rank of x (the largest x is assigned the rank 1) and M, β are constants. Therefore, if we rank the cases in a sample from large to small, the log of the item and the log of the rank should be linearly related (with a slope of β) for the Pareto hypothesis to be acceptable.

In the observed items it is not the case. A clear downward concavity of the distribution was observed for all items. This departure is very significant. Firms occupying the middle of the rank are more than twice as large as that predicted by the Pareto distribution. Figure 3 (page 10) shows an example of the shape raw accounting numbers assume when their logs are plotted against the logs of their rank. Cases follow much more closely the lognormal deviate (the dashed line) than a straight line. Ijiri and Simon [20] (1977) also report the same concavity for Sales and the Number of Employees in US firms.

The Pareto process and the growth of firms: In some literature it is usual to link the growth of firms with a Pareto process (see for example Steindl [28], 1965). If the growth of firms is Pareto-like, we should observe cross-sections of raw accounting numbers obeying the Pareto hypothesis. Since this is not the case, it seems as if our results contradict this belief.

The models of growth used to justify a Pareto distribution of firm sizes are inspired by the Gibrat law [17]. The Gibrat law leads to a whole class of skewed distributions depending on the conditions imposed on the growth process. As we recall in chapter 3, the most immediate outcome of the Gibrat law is lognormality. Lognormality, however, is too simple and general an outcome. It requires a random walk as the growth rate of firms. The literature concerning the growth of firms considers Pareto processes instead of the lognormal ones because of the scarcity of assumptions the lognormal hypothesis allows. When known influences like the serial correlation during growth, the effect of mergers and acquisitions and the birth and death of firms are accounted for, the resulting cross-sectional distribution should be a Pareto or a Yule one, not the lognormal.

1.5 Assessing the Importance of Multiple Tests

In this section we discuss the meaning and importance of W and corresponding P values obtained from applying the Shapiro-Wilk test to a large number of different samples. The Gaussian distribution is the result of many independent random causes. It would be interesting to compare our set of P values with the Gaussian distribution in order to measure in what extent its variability can be considered as caused by many independent random events as those influencing mechanisms of sampling.

A set of P values such as the one obtained from the repetition of a significance test for many samples cannot be directly compared with the Gaussian distribution because probabilities or any relative frequencies are bounded by 0 and 1. However, it is possible to transform probabilities so that the resulting variable is Gaussian for random normal events. We used a simple logit transformation to link relative frequencies with Gaussian deviates. For a given P, $[0 \leq P \leq 1]$, we computed

$$Logit\ P = \log \frac{P}{1-P}$$

$Logit\ P$ (also known as log-odds) ranges from $-\infty$ to $+\infty$ and is approximately normal for random normal events. A value of $Logit\ P = 0$ is the expected or central one. Negative logits mean $P < 0.5$ while the positive ones are obtained for $P > 0.5$. $Logit\ P$ can boldly be taken as the number of standard deviations for normal distributions.

Firstly we examined the distribution of the log-odds obtained after performing the $910 + 350$ tests described above. In order to do this we had to exclude the twenty bad tests having $P < 0.001$, because they would yield $-\infty$ for the number of significant digits we were working with. The resulting distribution had a mean value of -0.25 (equivalent $P = 0.49$) and a standard deviation of 0.85. Both skewness and kurtosis were very small. The aspect of the frequency distribution was nearly Gaussian. Next, we simulated an equal number of samples drawn from a strictly Gaussian population. These samples had the same number of cases as found in each performed test. Simulated normal deviates yielded a mean value of 0.04 (equivalent $P = 0.53$), and a standard deviation of 0.8. 3.6% of the simulated tests yielded $P < 0.05$. Notice that supposing normality, 95% of them would have to fall inside an interval of $\{-1.6, +1.6\}$ logits, which is $P = \{0.025, 0.975\}$, as they did.

The fact that when performing multiple tests some of them are expected to exhibit $P < .05$ even when the samples were drawn from a Gaussian population, means nothing wrong with the Shapiro-Wilk test itself. When many samples are drawn from a perfectly Gaussian population, it is likely that some of them will be far from normality in some degree due to the random nature of sampling.

Figure 4 shows the frequency distributions of both real and simulated tests. Some interesting conclusions arise from comparing them.

- When many tests are performed we can expect at least 3.6% of them to show significant departures from normality even when the population from which the samples were drawn is strictly Gaussian. In our case, we put aside 20 tests. The proportion of cases having $0.001 < P < 0.05$ is now 11.3%. It seems as if only something like 7.7% of those $(11.3 - 3.6)$ should be considered as real, unexpected, departures from two-parametric lognormality.

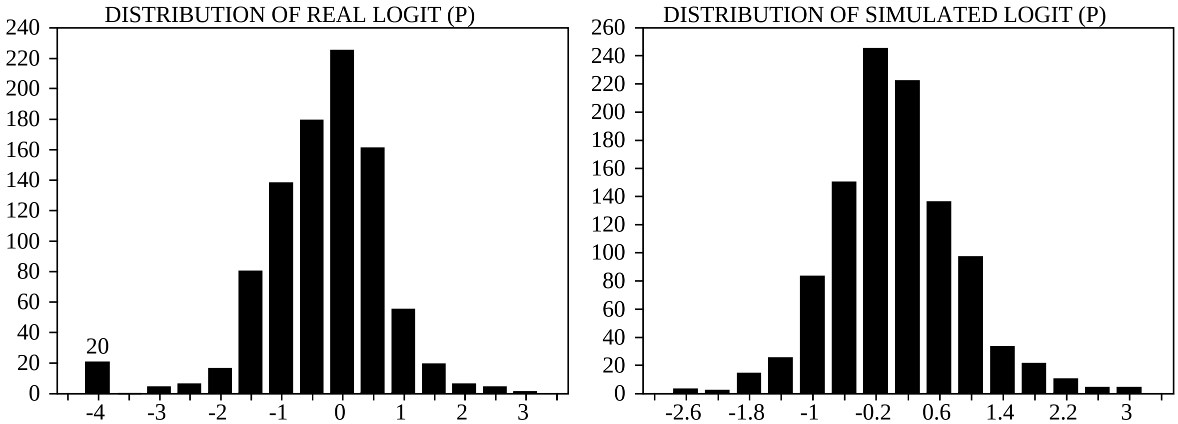

Figure 4: The frequency distributions for real (left) and simulated (right) *Logit P*. The cluster of the 20 "bad samples" has been added to the distribution of real *Logit P* with an arbitrary, small, value.

- The second conclusion is induced by the similarity of distributions of real and simulated results. Normality means random, independent, causes. Now, if some hazardous sampling can introduce a standard deviation of 0.8 in an otherwise perfectly Gaussian collection of samples, the standard deviation of the real tests, also 0.8, should be assigned to an hazardous mechanism of sampling and not to any particular cause.

If we take the difference between estimated expected values as the sole factor affecting the lognormality of items we can say that, while the expected probability associated with the Shapiro-Wilk W is, for a normal population, 0.53, it becomes 0.49 in the case of items. Four to one hundred odds is what separates items from a two-parametric lognormal mechanism.

We noticed in previous sections that twenty samples drawn from one industry at a time yielded a very small significance of W, denoting severe departures from lognormality. These bad samples didn't improve with the three-parametric transformation, and their Shapiro-Wilk's W significance were so small that it was impossible to apply logits. Such set of bad samples behave differently from the other ones. The significance obtained from all other tests form, in logit space, a nearly Gaussian distribution. Bad samples don't fit well in such a distribution. They are more numerous than expected and they form a cluster sticking out below the lower values of *Logit P* (See figure 4, page 12, on the left).

1.6 Discussion and Conclusions

Lognormality emerges as a general and stable feature of cross-sectional samples of raw accounting numbers. Not only the large samples containing many industries are lognormal. Each individual industry is lognormal too. Not only positive stocks are lognormal. Items with different origins, flows as well as stocks, are lognormal. This fact suggests that the mechanism determining lognormality in raw data has little to do with the way particular items are generated inside the firm.

The number of samples tested by this study is much larger than the existing ones. To what extent can the strong regularities described here be extrapolated to different samples? Our study clearly excludes firms which are too small to be collected into the Micro-EXSTAT database. In what concerns other industries, given the lognormality of samples obtained by gathering all groups together, lognormality is likely to be found also in industries not contemplated in this study. In fact, this overall homogeneity is the one expected when sampling firms at random obeying a single condition, the selected firms being industries.

We found that some samples were three-parametric. Two-parametric lognormality only contemplates situations in which the smallest cases approach zero. In raw accounting numbers, as in other lognormal variables, thresholds are expected. They have a role in explaining non-proportionality in ratios, as we shall see in section 4.3. We also noticed that three-parametric lognormality is scarce. It emerges in some years but not in others, and it relates to industries rather than to items.

Lognormality and cross-section: Hitherto no strong regularity was found when examining accounting data. This is because only ratios were tested. Ratios are exposed to the interaction between their components. This makes them different from one another. There is no reason to expect such a variety in the distribution of raw numbers.

Lognormality isn't difficult to explain. Any stochastic accumulation, that is, any growth proportional to the size already attained, leads to cross-sections that are lognormal or belong to the same class of skewed distributions. Notice that, in order to explain lognormality in accounting numbers, we don't need to accept a stochastic accumulation as the mechanism underlying every item. It is enough to picture the growth of the firm as a stochastic accumulation. In our data, all sorts of items are lognormal. Net Worth isn't particularly less lognormal than Sales despite not being a "sum of similar transactions whose sign remains the same" (McLeay [23] 1986, p. 79). Both Sales and Net Worth are lognormal, seemingly because the growth of the firm as a whole is itself an accumulation. Cross-sections of items reflect, on average, a given proportion of size.

Notice also that lognormality expresses an expected proportionality of random effects, not just a strict proportionality, as ratios do. The finding of lognormality in raw data is promising, not only because it answers many questions, including, to some extent, why ratios are used in financial statement analysis. It is also promising in that it opens up the possibility of improving ratios.

Chapter 2

Outliers, Heteroscedasticity and Trimming

What are the immediate conclusions to extract from the lognormal nature of the observed items? In this chapter we point out that the sole consideration of lognormality is enough to account for the persistent emergence of outliers referred to in the literature. The also mentioned heteroscedasticity of residuals is then discussed. We show that Least-Squares modelling of lognormal data isn't adequate. Other models ought to be developed. Specifically, we study the consequences of using ordinary or weighted regressions. Finally, the usefulness of trimming lognormal tails is discussed.

The meaning of the log transformation: In this study we often use the term "space" with a qualifier. For example, we refer to the rotated space or to the log space. Our goal is to emphasize the fact that a given set of variables has been jointly transformed in a well known and consistent way, thus defining a formal system.

The accounting literature is cautious about transformations. They are seen as a means of massaging data. For example, Eisenbeis [11] (1977) has been frequently quoted as saying that

> In the case of the log transformation there is also an implicit assumption being accepted where such a transformation is employed. That is, the transformed variables give less weight to equal percentage changes in a variable where the values are large than when they are smaller. If, for example, the variable being transformed was firm size the implications would be that one does not believe that there is as much difference between a \$1 billion and a \$2 billion size firms as there is between a \$1 million and a \$2 million size firms. The percentage difference in the log will be greater in the latter than in the former case.

This is not so. The log transformation has a precise meaning and shouldn't be considered as a mere manipulation of values to make them more tractable. Firstly, notice that

Log of one million	=	6	Log of one billion	=	9
Log of two millions	=	6.301	Log of two billions	=	9.301
The difference is		0.301	The difference is		0.301

Logs yield similar differences whenever the ranges are proportional, that is, when they are similar except for scale. In log space, a difference from one million to two millions is as impressive as a difference

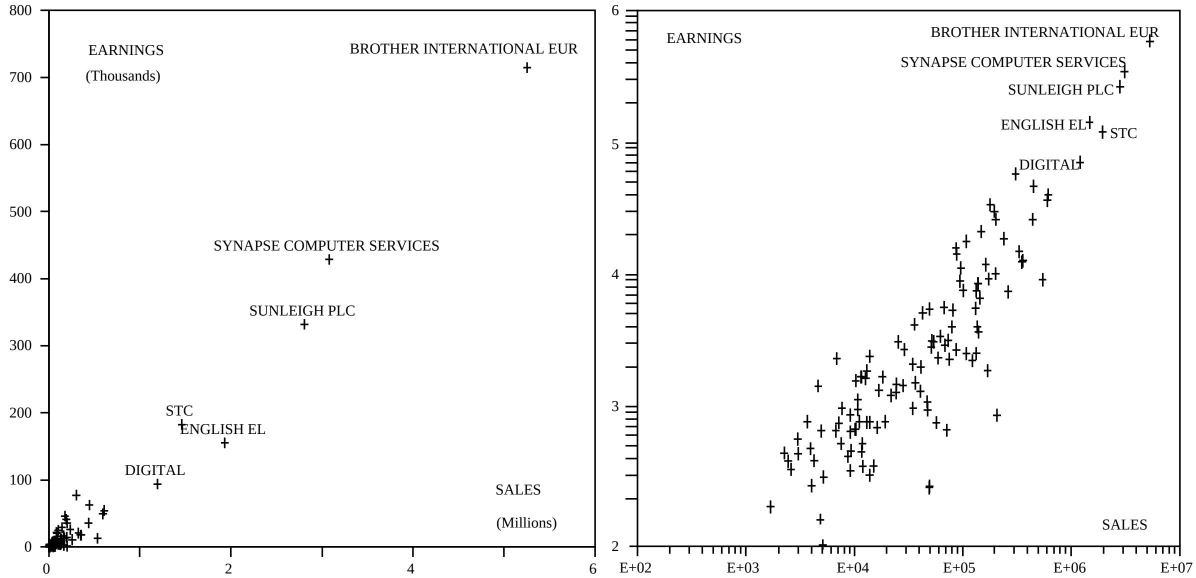

Figure 5: The awkward aspect of a bivariate lognormal distribution (left) and the homogeneous one of the same data in log space (right). Electronics, 1986.

from one billion to two billions. Of course, we should avoid applying proportions to log distances. It would be as if we were building proportions of proportions. This is Eisenbeis's pitfall.

The meaning of the log standard deviation: In log space, distances from any case to the mean are no longer real distances: A log displacement of 0.4, — a usual value for the log standard deviation observed in industries — means that the displaced case has been multiplied or divided by 2.5, the anti-logarithm of 0.4. Hence, we may say that a log standard deviation of 0.4 is equivalent to a multiplicative one of 2.5. Or that each unit of standard deviation in log space measures a scatter of 150%.

An immediate consequence of lognormality is that the coefficient of variation (the standard deviation expressed as a proportion of the expected value of the variable) is constant. Hence, in lognormal distribuions, the expected variance is not constant (homoscedastic): It grows with the square of the variable.

2.1 Outliers

The presence of outliers in ratios is consistent with the lognormality of items. For standard deviations like those found in accounting raw data, the lognormal distribution is severely skewed. It is this skewness that has been taken as outliers by the literature. In fact, strongly skewed distributions have long tails towards the positive values of observations: Many cases are concentrated in a small region and a few of them spread out along a large range. Figure 5 on the left (page 21) is a bivariate example. Apart from the biggest companies, the remaining ones (about 140) occupy a small region at the bottom left of the plot. It is easy to take the few extreme values as outliers. As Snedecor and Cochran put it, ([27], 1965, pp. 281, 9^{th} ed.) "the apparent outliers may reflect distributions of the observations that are skew or have long tails". This figure is also an example of the adequacy of the log space to raw accounting data. On the scatterplot on the left, hardly anything can be sorted out. When drawing the same plot in log space (on the right), each case becomes separable. A small non-linear relation between the two components is now visible. This non-linearity turns out to be important to understand non-proportionality in ratios.

Along this study we shall see that other pieces of evidence, hitherto hidden from direct observation, become visible in log space.

Lognormality and a very small scatter: Notice that the above description of the characteristics of the lognormal distribution shouldn't be taken as a general rule. Perfectly lognormal samples can have small skewness when their standard deviations are very small too. It is generally accepted that coefficients of variation smaller than 0.25 denote distributions which can be approached by the Gaussian. Such exceptional cases cannot explain why a few ratios are Gaussian: Ratios seldom exhibit a coefficient of variation small enough to comply with the above condition.

2.2 Heteroscedasticity

There is a well known claim for regressions to be used instead of ratios because of non-proportionality between ratio components (Barnes [2], 1982). On the whole, the discussion resulting from this claim was very revealing. It drew the attention of researchers for the use of raw numbers instead of just ratios, and it introduced an aspiration towards more accurate models. However, given the lognormality of items, the use of regressions is inadequate. In fact, lognormality implies leverage cases for the same reason it implies outliers. Leverage cases are likely to be influential, thus distorting regressions.

Statistical models suffer distortions if one or two input vectors are influential. When an influential case is excluded, the fitted parameters are significantly different from when it is present. This happens because the quadratic nature of Least-Squares minimization precludes large residuals.

In regressions using raw data as input variables, a large firm is a leverage case, likely to become influential, just by being large. In fact, when a lognormal distribution is taken as Gaussian, the outliers are always in its tail — the largest firms. In the literature this problem is referred to as the non-acceptable heteroscedasticity of data. But there are many degrees of heteroscedasticity. As stressed by Berry and Nix [4] (1991), in order to cope with heteroscedasticity, knowledge about its nature is required. Recipes adequate for one particular form don't work in different cases. Next we comment on one of such recipes widely used in accounting research.

Weighted Least-Squares: A weighted regression uses

$$\frac{y_j}{x_j} = \frac{a}{x_j} + b + \frac{\varepsilon_j}{x_j} \qquad \text{instead of} \qquad y_j = a + b \times x_j + \varepsilon_j .$$

For a variance increasing with x this procedure stabilizes it. But in lognormal deviates it's the standard deviation, not the variance, which increases with x in average. The variance grows with the square of x. Figure 6 on page 23 shows the meaning of this distinction. The scatterplot above is a bivariate relation ideal for a weighted Least-Squares transformation. In fact, this sample has been obtained from a perfectly homoscedastic sample, by applying the inverse of a weighting transformation. Below, the same homoscedastic sample, but after applying the anti-logarithmic transformation. The aspect of both sets is typical of data requiring weighting (above) or logs (below).

The Cook Distance: As an example of the correlation between influential cases and the size of the firm, we selected the sample displayed in figure 5, page 21 (Electronics industry, 1986, EBIT with

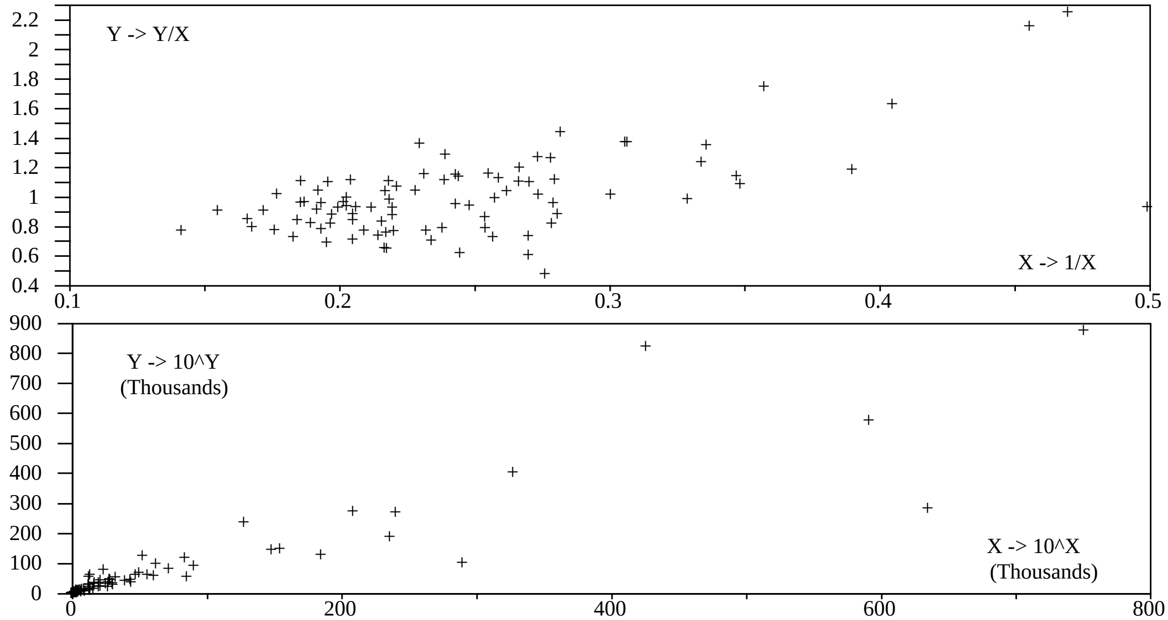

Figure 6: The typical aspect of bivariate distributions requiring weighted Least-Squares (above) and Log transformation (below).

Ordinary Least-Squares (OLS)			Weighted Least-Squares (WLS)			OLS in log space (LOG)		
Name	Rank	Cook D.	Name	Rank	Cook D.	Name	Rank	Cook D.
G. E.	1	12.18	TELFORD	129	7.890	NATIONAL	17	0.077
STC	4	1.485	MISYS	142	0.166	G. E.	1	0.073
IBM UK	2	0.985	M.M.T.	145	0.152	POLYTEC.	126	0.057
ENG. EL.	3	0.250	FORWARD	139	0.079	IBM UK	2	0.053
STC C.	6	0.213	KLARK-TK	137	0.057	BELL & H.	59	0.039
DIGITAL	7	0.062	HEADLAND	135	0.053	CASIO	55	0.039
AMSTRAD	14	0.033	AMS IND.	136	0.048	AMS IND.	136	0.036
UNISYS	27	0.013	SOUNDTR.	138	0.044	ZYGAL	125	0.035
FERRANTI	10	0.012	AVESCO	131	0.039	ENG. EL.	3	0.034

Table 9: The largest Cook distances when Sales was used to explain EBIT in three different regressions: OLS, WLS and LOG. Electronics, 1986.

Sales). Three different regressions — OLS, WLS and LOG (a regression in log space), were performed. Sales was the input and EBIT the outcome. Then, we compared the Cook distances observed in each one of them. The Cook distance (Cook [9], 1977) measures the effective degree in which every case in the sample commands the whole fit. Tolerable Cook distances should not exceed 1. Values larger than 1 mean a fit monopolized by the case in which it occurs. Table 9 shows the firms which were traced as most influential in each regression. The column labeled "rank" signals the ranking of each firm according to size. In this column, 1 means the largest firm, and so on.

OLS has two firms which are influential. WLS has one. When using OLS, the influential firms tend to be the largest ones. In the case of WLS, they tend to be the smallest ones. The Cook distances in log space are far below those of OLS or WLS. Notice that, when using WLS instead of OLS, some improvements are observed. Weighting makes the standard deviation of the variables smaller, and the skewness, in lognormal distributions, depends on the standard deviation.

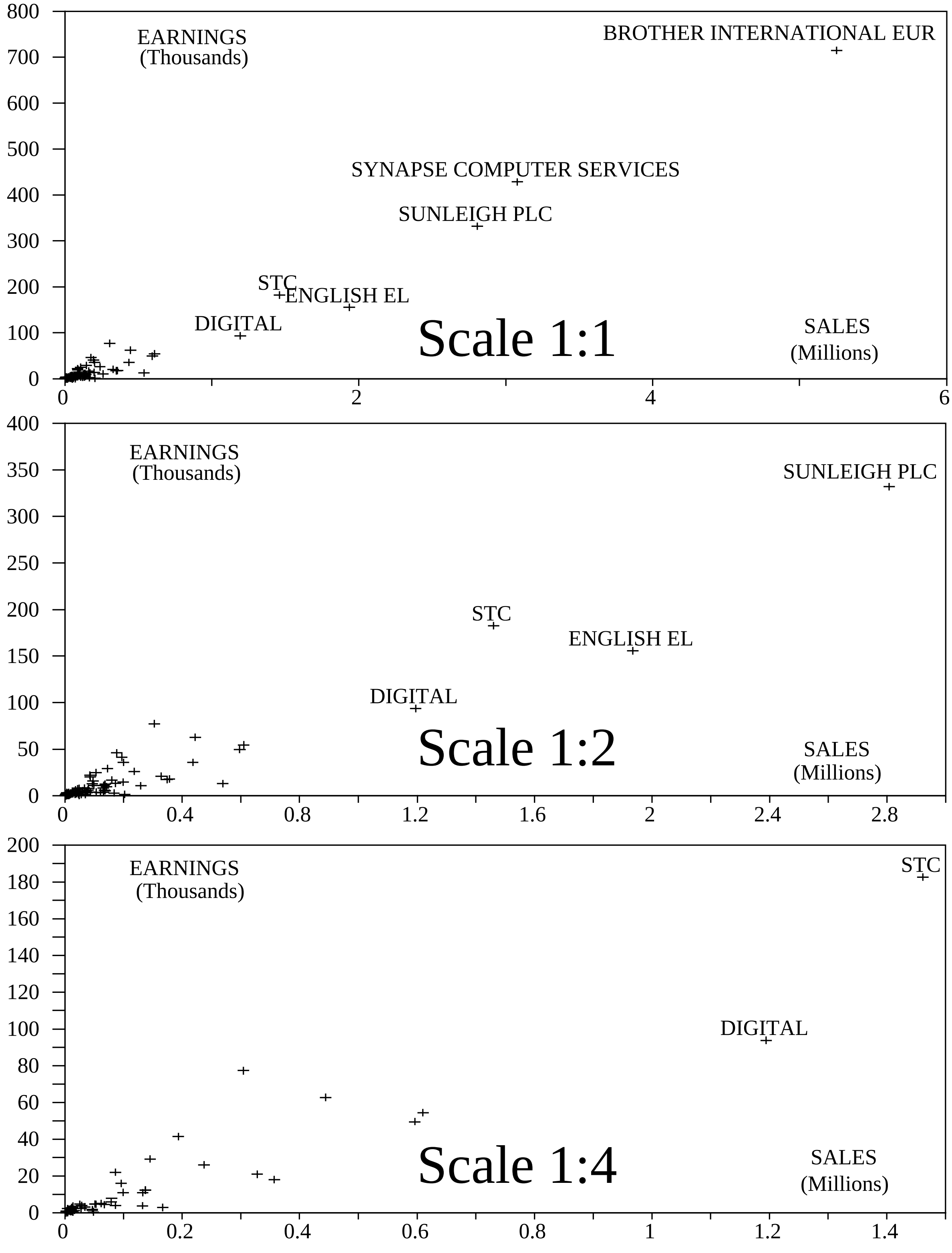

Figure 7: When trimming outliers caused by a lognormal scatter, new outliers emerge across scales. Electronics, 1986.

2.3 Trimming

Another point the lognormal nature of items elucidates is why it seems so unfruitful to trim outliers.

Lognormal multivariate distributions exhibit self-similarity of features across scales. Any pattern which holds for billions also holds for thousands. The shape like a “<”, typical of a lognormal bivariate scatter of correlated variables, along with the corresponding gradient in the density of cases, is continuously generating influential cases across scales. As a consequence, there is little point in excluding large firms from the sample to obtain a more homogeneous set. If we exclude the largest cases, new ones will emerge as outliers since the phenomenon commanding the emergence of influential cases holds in different scales.

In the above example (table 9 on page 25, the OLS case), if we measure the Cook distance associated with each case after excluding the two influential firms (G.E. and STC), we get three new cases with a non-acceptable weight in the regression: SUNLEIGH PLC (with a Cook distance of 1.6), ENGLISH ELECTRIC (1.9) and BROTHER INTERNATIONAL (19.8). The new situation is worse than before. If we exclude also these three firms, SYNAPSE COMPUTER SERVICES emerges with a new Cook distance of 80.5. And excluding also this firm, MISYS PLC appears as having a new Cook distance of 0.8, much higher than the highest values observed in LOG.

Notice that DIGITAL, despite being amongst the largest firms and therefore a leverage case, didn't become influential. This is because leverage and influence are different concepts. An influential case must be a leverage case. But it is possible for a leverage case not to become influential.

2.4 Conclusion

The variability generated by proportional mechanisms is well known and documented. For example, Snedecor and Cochran [27] (9^{th} edition, 1989, pp. 290) observe:

> “Logarithms are used to stabilize the standard deviation if it varies directly with the mean, that is, if the coefficient of variation is constant. When the effects are proportional rather than additive, the log transformation brings about both additivity of effects and equality of variance.”

The log transformation solves at once two problems that the literature considers separately: The non-homogeneous variance and the emergence of outliers.

Regressions shouldn't be used to model relations between lognormal variables. Lognormal distributions generate large residuals which monopolize the minimization of square errors. The results are then dependent on one or two influential cases.

Weighted Least-Squares isn't an appropriate technique to deal with the above problem. It simply transfers the influence from the largest to the smallest cases in the sample. The heteroscedasticity would not vanish in any of these cases. The log transformation is adequate, in a first approximation, to render residuals additive. But appropriate models ought to be developed that fully explore the existence of non-proportional and non-linear relations in items.

The trimming of outliers becomes a useless exercise for lognormal data. The shape like a “<”, typical of bivariate proportional relations, will not change across scales. It will generate more and more outliers if successive trimming is to be attempted.

Chapter 3

The Overall Effect of Size

In this chapter we study the co-variance matrices formed, in log space, with items from the same report. We show that raw numbers share most of their variability. They are well described as a unique effect with some particular randomness superimposed. We also discuss the statistical behaviour of items obtained by subtraction of other items. We explain why positive differences maintain lognormality and we suggest procedures to use, in log space, variables having negative cases.

The observed data is the same as in chapter 1. It contains 14 industrial groups and 18 items. This sampling is replicated during a period of five years (1983-1987). Tables 2 on page 14, and 1 on page 14 describe its contents.

3.1 Co-Variance Matrices in Log Space

Inside industries, the cross-sectional variances and co-variances of log items from the same report are similar and lie above a common ground-value. Also, for samples of significant size, no negative or zero co-variances can be observed. Table 10 shows four co-variance matrices (usually referred to as Σ). They describe the features we highlight. In appendix (figure 29 on page 98) other matrices are displayed.

Σ matrices such those of table 10 are only possible if there is a preeminent source of common variability. It seems as if log items were just the same variable with a bit of particular randomness superimposed. This is consistent with the fact that, in cross-section, the first and most important

	S	NW	W	I	CA	FA	EBIT	S	NW	W	I	CA	FA	EBIT
S	0.35							0.21						
NW	0.32	0.39				Paper and Pack.		0.22	0.26				Leather	
W	0.35	0.36	0.41					0.19	0.21	0.18				
I	0.33	0.34	0.38	0.40				0.19	0.21	0.18	0.19			
CA	0.33	0.32	0.35	0.35	0.34			0.21	0.24	0.20	0.20	0.23		
FA	0.32	0.36	0.37	0.34	0.33	0.40		0.19	0.21	0.18	0.17	0.19	0.20	
EBIT	0.35	0.35	0.36	0.35	0.34	0.35	0.41	0.26	0.29	0.24	0.23	0.27	0.24	0.28
S	0.17							0.55						
NW	0.17	0.20				Clothing		0.57	0.73				Food Manuf.	
W	0.18	0.18	0.23					0.55	0.64	0.64				
I	0.19	0.20	0.20	0.23				0.59	0.68	0.63	0.73			
CA	0.17	0.18	0.18	0.20	0.19			0.56	0.63	0.58	0.65	0.61		
FA	0.17	0.19	0.21	0.20	0.17	0.26		0.57	0.69	0.65	0.65	0.61	0.71	
EBIT	0.17	0.17	0.18	0.18	0.17	0.19	0.31	0.59	0.71	0.64	0.70	0.66	0.69	0.73

Table 10: Four Co-variance matrices denoting the homogeneity of variance and co-variance inside industrial groups. Data from 1983.

source of variability impinging upon different items is size. Only after accounting for the effect of size is it possible to assess the variability specific to items.

In order to provide a more systematic evidence on the existence of this strong source of common variability, we created a variate, s, supposed to capture it. This variate can be built in several ways. Ours will be explained in section 6.1. Using s as the predictor, we then formed regressions in log space. Each log item was explained by s ($\log x_j = a + b \times \log s_j + \varepsilon_j$). Finally, we observed the obtained R^2.

Results: Tables 30 and 31 in appendix (pages 96 and 97) contain the detailed results of this experiment, by industry and by year. Table 11 on page 29 shows the results for all industries together. We also present a condensed table, the 12^{th}, in which the R^2 are the mean of five years, by industry and item. This table is sorted in ascending order of its marginal content.

The results are similar during the usual period of five years. The R^2 range from 97% to 75% for most of the items. Earnings and Working Capital are below the usual and Long Term Debt has an R^2 ranging from 30% to 80%, smaller than any other item. However, even in the case of Debt it isn't possible to accept independence from size as an R^2 of 30% means a significant correlation of 0.55. The conclusion is that it is possible to create a unique variable, s in this case, able to explain most of the randomness of our set of log items. The results would be similar to these if we would use, instead of s, another item, say, $\log TA$ or $\log NW$.

It could be argued that these large R^2 stem from using items both for building s and for explaining them in terms of s. The observed co-variance would reflect the effect of the same item in both sides of the regresion. We shall see in section 6.1 that items like S, NW, W, D, CA, and CL, were used in the building of our proxy for size. However, this fact will not change the results. The R^2 of items which weren't used for building s, like I, FA, C, FL, and others, aren't significantly smaller than the R^2 of those which were. The large R^2 obtained can only be explained by an effect common to all items.

The slopes of regressions in which log items are explained by s, remain close to 1 (see table 11). This stems from applying log transformations. In log space, any scaling becomes a translation. But translations can be accounted for by the constant term of regressions. Hence, the main source of variability in log items is modelled by a simple mean adjustment. Departures from 1 in the value of the slopes are explainable by using OLS instead of errors-in-variables modelling, and by the presence of significant thresholds.

In the above experiment, the negative-valued cases were excluded from the sample. However, we also observed Σ matrices formed from groups of firms sharing the same problem. For example, the matrices displayed in figure 8 on page 29, belong to the Electronics industry (1986). On the left, only firms with negative Working Capital were selected. On the right, only firms with negative EBIT.

These patterns are not similar to one another, denoting the different nature of their underlying mechanisms. Samples with negative Working Capital generate more variability than those with negative EBIT. Notice that the number of cases is smaller than the desirable. We obtained a statistic, Σ, which has more parameters in it than the number of cases in the sample. Such an analysis is case-dependent.

Comparing simulated and real Σ: Simulation can be carried out in order to observe, avoiding the burden of analytical developments, the multivariate pattern of Σ matrices when they are calculated from the logs of absolute values of negative cases. This allows us to understand which features of the above matrices are due to a particular behaviour of firms — liquidity or profitability problems for example — and which are due to the mechanism of subtracting two lognormal variables.

Item	1983		1984		1985		1986		1987	
	slope	R^2	slope	R^2	slope	R^2	slope	R^2	slope	R^2
S	1.006	0.91	1.006	0.89	0.994	0.90	0.991	0.90	1.004	0.89
W	1.027	0.91	1.046	0.91	1.040	0.92	1.033	0.93	1.036	0.93
NW	1.009	0.89	1.010	0.88	1.008	0.88	1.008	0.90	0.991	0.89
I	1.047	0.88	1.054	0.88	1.029	0.86	1.055	0.86	1.043	0.82
D	0.991	0.91	0.972	0.91	0.977	0.91	0.974	0.91	0.985	0.91
C	1.008	0.90	1.015	0.92	1.022	0.92	1.016	0.91	1.017	0.91
CA	0.996	0.94	0.985	0.95	0.995	0.94	0.992	0.94	0.976	0.94
FA	1.075	0.81	1.061	0.82	1.099	0.82	1.103	0.83	1.100	0.82
CL	0.981	0.92	0.986	0.93	0.984	0.93	0.992	0.94	0.984	0.93
FL	1.047	0.84	1.052	0.85	1.033	0.87	1.046	0.85	1.035	0.86
EBIT	1.036	0.78	1.032	0.81	1.024	0.82	1.022	0.83	1.008	0.83
N	0.987	0.84	1.004	0.85	1.010	0.86	1.009	0.87	1.020	0.86
DEBT	1.157	0.62	1.096	0.61	1.100	0.60	1.107	0.56	1.142	0.56
EX	1.000	0.86	1.006	0.84	1.001	0.84	0.990	0.84	0.997	0.83
WC	0.996	0.74	0.976	0.80	0.967	0.75	0.978	0.74	0.934	0.76

Table 11: The slopes and explained variability (R^2) obtained when $\log s$, an estimated size, is used to explain several log items. All groups together.

	DB	WC	EB	FA	FL	NW	I	D	C	CL	S	W	CA	mean
METAL	0.40	0.81	0.80	0.59	0.78	0.77	0.81	0.83	0.79	0.78	0.65	0.81	0.90	0.75
CLOTH	0.32	0.56	0.54	0.76	0.67	0.84	0.89	0.79	0.86	0.85	0.92	0.87	0.92	0.75
TOOLS	0.31	0.73	0.70	0.83	0.69	0.78	0.88	0.82	0.92	0.88	0.94	0.95	0.95	0.80
WOOL	0.41	0.78	0.80	0.81	0.86	0.85	0.77	0.88	0.85	0.82	0.91	0.86	0.91	0.81
PLANT	0.57	0.67	0.67	0.82	0.77	0.80	0.88	0.93	0.90	0.93	0.96	0.96	0.95	0.83
PAPER	0.46	0.71	0.84	0.77	0.86	0.83	0.88	0.86	0.91	0.94	0.91	0.92	0.96	0.83
ELECT	0.49	0.71	0.79	0.74	0.84	0.85	0.90	0.92	0.95	0.95	0.93	0.93	0.97	0.84
CHEM	0.62	0.76	0.84	0.84	0.84	0.89	0.80	0.88	0.90	0.92	0.87	0.95	0.92	0.85
ELTN	0.53	0.84	0.80	0.81	0.86	0.92	0.79	0.93	0.90	0.93	0.93	0.90	0.93	0.85
BUILD	0.62	0.74	0.82	0.87	0.86	0.88	0.92	0.96	0.90	0.95	0.97	0.97	0.97	0.88
FOOD	0.73	0.77	0.87	0.86	0.90	0.89	0.90	0.93	0.93	0.94	0.91	0.94	0.94	0.88
MOTOR	0.76	0.80	0.82	0.92	0.88	0.92	0.94	0.96	0.91	0.95	0.96	0.96	0.90	0.90
MISC	0.77	0.78	0.90	0.81	0.92	0.93	0.94	0.94	0.94	0.93	0.96	0.90	0.95	0.90
LEATH	0.68	0.90	0.93	0.87	0.94	0.95	0.97	0.85	0.94	0.96	0.98	0.93	0.97	0.91
mean	0.55	0.75	0.79	0.81	0.83	0.86	0.88	0.89	0.90	0.91	0.91	0.92	0.94	

Table 12: The R^2 (mean of the five years observed) of regressions in which s, an estimated size, explains 13 items. Rows are industries and columns are items.

S	W	I	D	C	CL	FA
0.63					13 cases	
0.41	0.33					
0.54	0.39	0.61				
0.53	0.38	0.51	0.49			
0.47	0.33	0.45	0.45	0.48		
0.48	0.35	0.48	0.46	0.45	0.47	
0.60	0.41	0.53	0.54	0.49	0.48	0.67

S	W	I	D	C	CL	FA
0.32					23 cases	
0.24	0.28					
0.32	0.28	0.47				
0.28	0.26	0.32	0.30			
0.23	0.22	0.28	0.24	0.23		
0.25	0.23	0.28	0.26	0.23	0.26	
0.24	0.26	0.25	0.26	0.23	0.24	0.33

Figure 8: Two Σ matrices obtained from the same industry and year. On the left, cases with negative Working Capital. On the right, cases exhibiting negative EBIT.

S	CA	FA	CL	
0.510				S
0.478	0.504			CA
0.490	0.480	0.599		FA
0.462	0.470	0.467	0.460	CL

Figure 9: Σ matrix used for simulating Working Capital, 1986, Electronics.

	S	FA	CL	CA	WC		S	FA	CL	CA	WC
S	0.516		W. Cap. > 0				0.462		W. Cap. < 0		
FA	0.505	0.627					0.416	0.475			
CL	0.464	0.477	0.459				0.419	0.401	0.423		
CA	0.475	0.484	0.465	0.489			0.422	0.400	0.424	0.429	
WC	0.494	0.492	0.474	0.532	0.672		0.399	0.407	0.416	0.392	0.680

Figure 10: Σ matrix from the simulated Working Capital. Electronics, 1986.

We used the same group, Electronics 1986, and starting conditions similar to those found in positive cases. Such conditions were:

Item	Mean	St. Deviation	Item	Mean	St. Deviation
Current Assets	4.213	0.723	Fixed Assets	3.757	0.779
Sales	4.577	0.720	Current Liabilities	4.005	0.680

In figure 9 we display the Σ matrix used to introduce in the simulation the co-variance of items. After generating 2,000 cases with multivariate lognormality we obtained 185 with negative Working Capital. Both samples were lognormal but the positive one had positive, though small, skewness. The negative one had negative, also small, skewness and a larger kurtosis. The resulting co-variance matrices for both samples are displayed in figure 10. They aren't different from the usual. Hence, the differences observed in matrices obtained from real samples are probably caused by mechanisms of the firm.

We carried out simulations of other groups. The only feature recognizable as particular to simulated negative cases was its larger spread and the described skewness and kurtosis.

3.2 The Subtraction of Two Items

Many items from financial statements, namely those representing flows, are obtained by subtracting two other items. In this section we try to answer two questions related to the statistical characteristics of such items. These questions are:

- Why are they lognormal? In general, there is no reason why the subtraction of two lognormal variates should remain lognormal.
- If an item has negative cases the log transformation cannot be applied. This is the case for Earnings or Funds Flow. How can they be used in log space?

After explaining their lognormality we show that the problem of negative cases is not specific to the log transformation. We suggest alternative solutions enabling their use in log space.

Samples containing positive cases only: The lognormality of positive differences between two lognormal variables is a consequence of a strong correlation between them. In a subtraction, $z = y - x$, of two correlated items, cases in which y is large also have proportionally large x. And cases in which y is small are expected to have proportionally small x.

The extent to which z follows x and y is dependent on this correlation. For a pair of exactly synonymous items, z would be proportional to them. If s is the effect common to both x and y we can write in such an extreme case

$$\frac{y_j}{s_j} = R_y \quad \text{and} \quad \frac{x_j}{s_j} = R_x \quad \text{for any firm } j.$$

R_y and R_x are constants. Therefore $z_j/s_j = R_y - R_x$ for any j.

The more general case can be described by introducing in the above expression small f_j^y and f_j^x so as to reflect the variability particular to x and y:

$$\frac{y_j}{s_j} = R_y \times f_j^y \quad \text{and} \quad \frac{x_j}{s_j} = R_x \times f_j^x \quad \text{for any } j.$$

These f_j account for departures from a strict correlation with s in case j. If such departures are small, they should be near the unit. Considering $f^y = 1 - \eta^y$ and $f^x = 1 - \eta^x$ we can write

$$\begin{aligned} \frac{z_j}{s_j} &= R_y \times (1 - \eta_j^y) - R_x \times (1 - \eta_j^x) \\ &= R_y - R_x - d_j. \end{aligned}$$

The departure from proportionality is now isolated. It is $d_j = R_y \times \eta_j^y - R_x \times \eta_j^x$. Since the f_j are near the unit, the η_j are small. For values of R usual in ratios, d_j is a difference of two small values. Hence, z will be near lognormality for the same reasons x and y are lognormal.

Samples containing both positive and negative cases: In chapter 1 we showed that, when taken separately, both the samples of positive cases and those having negative ones, are lognormal. We can see why, by subtracting graphically two lognormal scatters correlated with size, as in figure 11. Given $WC = CA - CL$ we start with a scatterplot of CA on any measure of size and then we subtract CL from CA. The values of WC are obtained from CA by sliding them down a value which is CL. But the CL are proportional to size too. Hence, a large firm's CA is expected to be largely modified, and a small firm's CA is expected to change proportionally less.

Notice that the X-axis slices the bivariate distribution of WC with size. There are now two regions separated by the X-axis. But such regions preserve the proportional nature of the data. If the slicing were made along a line not passing through zero, or in less correlated items, the result would be a truncation. But whenever the slicing of a bivariate distribution of very correlated lognormal items is made along an axis of the distribution itself, the resulting two scatters will project their values in the Y-axis in a way that preserves lognormality. A subtraction of two items so that negative cases emerge, generates a juxtaposition of approximately lognormal distributions. One of them contains the positive cases. The other one, the mirror-image of the absolute value of the negative ones. The simulations we carried out corroborate this.

3.2.1 How to Model With Samples Having Negative Cases

What do we lose by taking separately one sample with positive cases and another one with the negative ones as we have done so far? At first, the existence of negative cases in a sample seems unsatisfactory for modelling purposes. By taking both sets separately, we break the continuity of the sample and lose the information describing the passing through the zero value. It would seem desirable to be able to work with the whole set of cases as a unique variable in the sample.

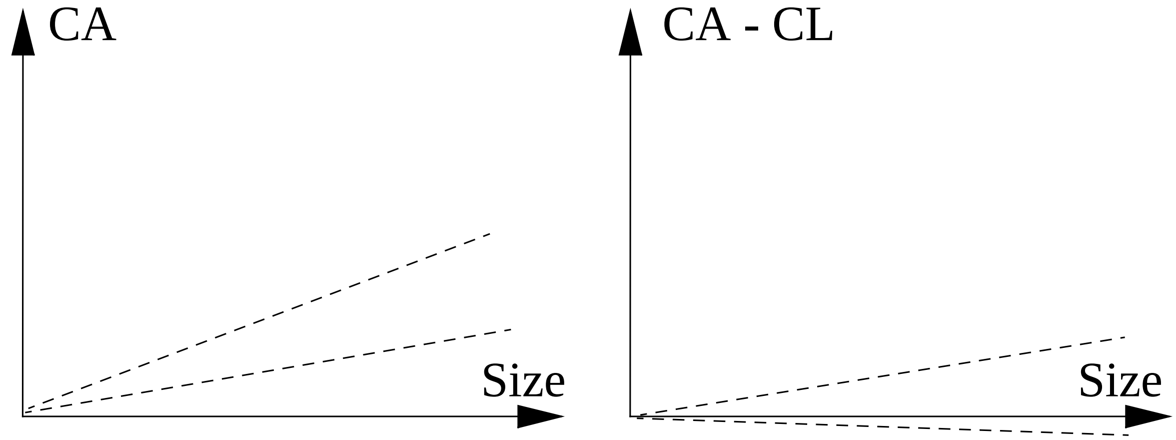

Figure 11: Schematic representation of the subtraction of two items yielding a new one with a few negative cases.

It is easy to recall important pieces of research in which the used ratios had values passing through zero. For example, Beaver's classic study on the importance of ratios for tracing firm's failure [3] shows how revealing a ratio of Cash Flow to Total Debt can be when sliding down from positive to negative values during an observed period. The consequence seems to be that ratios should be considered as a unity and taken as a whole. Breaking them into two samples, apparently damage part of their information content. According to this, the literature has devoted some effort to the assessment of the distribution of such ratios (see, for example, McLeay [23],[24] 1986). In fact, we lose nothing by using split samples in cross-sections. Beaver's ratios draw a trend during several time periods. One unique object — the average ratio for a group of firms — is observed during consecutive intervals. But cross-sections are not about one unique object. They capture the behaviour of many objects ideally at the same instant in time. The concern referred to above stems from picturing time-series and transposing it to cross-sections.

Ratios and the log transformation: The lack of continuity between positive and negative cases isn't a problem specific to the log transformation. Ratios face the same problem. Ratios — or more elaborated models like regressions — fail to model the same samples the log transformation is not apt to model. In fact, for the ratio y/x in which y is an item having both positive and negative cases, when we go along decreasing values of y and pass through zero, the corresponding evolution of x change in direction. It ceases decreasing and begins to increase. It is impossible to model such a sample using one ratio. For each x there are two possible y. That's why practitioners calculate standards by considering only positive values. They can't find a consistent standard for samples with both signs.

For example, consider the ratio Cash Flow to Total Debt. In a sample there are firms with large positive Cash Flow and large Debt. But there are also firms having negative Cash Flow and large Debt. When producing a ratio to explain the joint behaviour of these two items, the firms with negative Cash-Flow push the expected value of the ratio towards zero. The obtained estimation for Debt, given the Cash-Flow, is larger than it should be. The expected value of the ratio can even approach zero or become negative. When approaching zero, the amount of Debt predicted by the ratio rises to infinity. After passing through zero, the ratio predicts infinitely large negative Debt.

There is a breakage of continuity when passing through zero because each sample — positive Cash

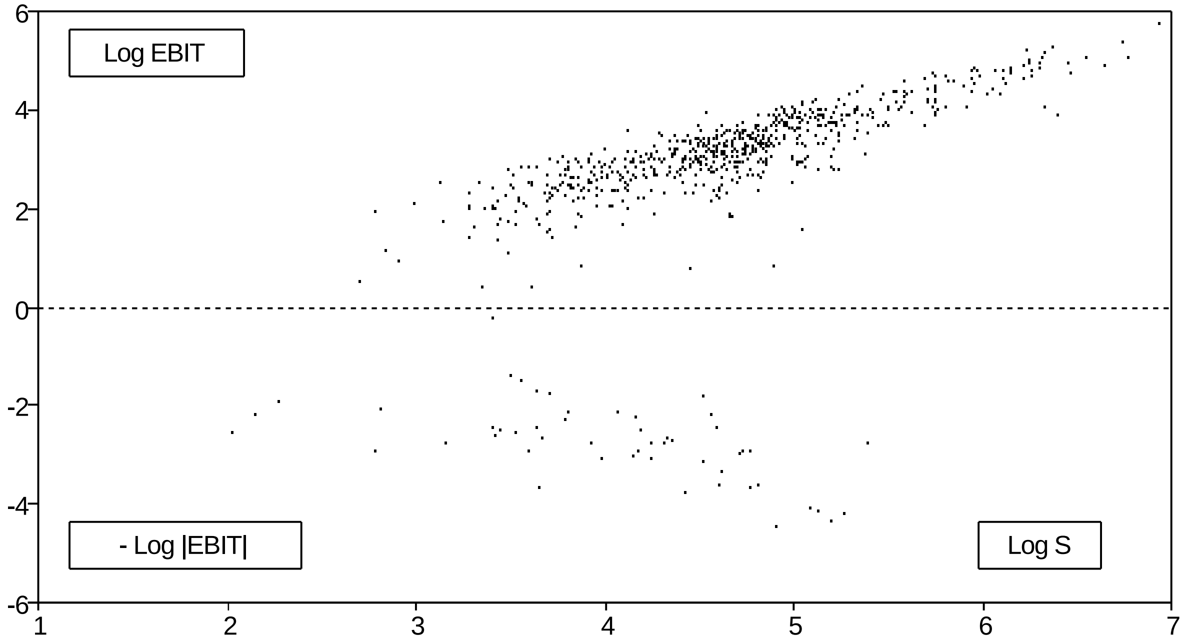

Figure 12: Scatterplot showing the effect of a log symmetrical transformation on the EBIT (Y-axis) to Sales (X-axis) relation. All industries, 1984.

Flow and the negative one — represents one group. They shouldn't be mixed up. Clearly, the problem is not in the logs or ratios. The problem is that two different groups cannot be modelled by the same parameter. In the case of ratios or logs, the algebra itself precludes one single model. Proportional co-movements cannot pass from one Cartesian quadrant to the other one except by going both together through the origin. Ratios, the same as logs, entail an assumption of proportionality. Ratios, an algebraic one. Logs, a differential one. Lev and Sunder [22] (1978) devote to this breakage of continuity a large comment.

This lack of continuity seems to make sense also on grounds of financial analysis. We recall the patterns of joint variability displayed in section 3.1 for samples reflecting problems of liquidity or profitability. They are very particular, contrasting with the general rules governing the positive ones. Negative cases reflect firm illness. On the contrary, the positive ones reflect a healthy state. In statistics, situations as those require a grouping variable.

3.2.2 Alternative Variables

Despite the above remarks, there are cases in which it is useful to mix up positive and negative cases. Here we suggest three possible solutions allowing their use in log space. Firstly, we can use a symmetrical transformation:

$$\begin{aligned} x &\mapsto \log(x), && \text{for } x > 0 \\ x &\mapsto -\log(|x|), && \text{for } x < 0 \end{aligned} \tag{1}$$

or the equivalent creation of a dummy variable. Such transformations correspond to the fact that negative cases are also lognormal and correlated with size. They are useful provided no attempt is made to fit a unique parameter to the transformed data. Figure 12 shows an example.

Secondly, if we scale one of the components of a difference so that negative cases cannot occur in the sample, we obtain a new item which is also a difference but has only positive cases. Scaling is equivalent to a translation in log space. Samples will not change their shape by introducing a scaling. Therefore, if we use $z' = y - x \times S$ instead of $z = y - x$ (S being a constant) we can work with this new variable knowing that the shape of z' is the same as the one of z, its standard deviation in log space didn't change and even the outliers are all there. Only the mean value has emerged a bit.

Finally, this problem is solved by using other items instead. The variability of Earnings can be brought to a model by Sales and Expenses. Working Capital can be substituted by Current Assets and Liabilities. For any item z, resulting from a subtraction of two positive items y and x, the pair $\{y, x\}$ obviously contains all the information of z, and a bit more.

The addition of constants: Indeed, there is one transformation which should not be applied for it severely distorts distributions. It consists of adding a large positive constant to all cases in the sample so that the negative ones become positive. This practice has been reported in a few studies.

Modelling Debt: The log transformation cannot be applied to items having cases with values of zero, like Debt. We avoid this problem by using, instead of $\log 0$, a very small number: $\log 1 = 0$. The use of $\log(x+1)$ instead of $\log x$ is recommended for such cases, in texts like Snedecor and Cochran [27]. This will be acceptable if, in the sample, there are no cases with values near zero. If a scaling of one million pounds is used instead of the usual scaling of one thousand, we are likely to find cases with values near zero, both positive and negative.

3.3 Summary

The important point this chapter outlines is the existence of a common source of variability in the logs of raw accounting numbers. In log space, items are the addition of two effects: The first one is preeminent, reflecting the relative size of firms. The second one, particular to each item, reflects its uniqueness. Earnings or Gross Funds From Operations are as correlated with size as Total Assets or Net Worth. The item showing a distinct behaviour is Debt. But even in this case size is present.

We also studied the problems posed by items having both positive and negative cases. We pointed out that, in cross-section, there is no continuity between the positive-valued cases and the negative-valued ones. We further suggested that negative-valued cases should be viewed as a different group since they represent specific situations.

Chapter 4

Validity and Extension of Ratios

What can be said about ratios and their validity, based solely on the findings of previous chapters? If raw numbers are Gaussian in log space, an observation x_j from report j could be explained as the expected value of the transformed variable, μ_x, plus a deviation, e_j^x. That is,

$$\log x_j = \mu_x + e_j^x. \tag{2}$$

In ordinary space, $\exp(e_j^x)$ reflects the number of times x_j is larger or smaller than the expected for its industry. For example, the expected value of the logs of Sales in the Food industry, was estimated as $\overline{\log x} = 4.9218$, (83,521 thousand pounds) in 1987. Then, we could say that G. F. LOVELL PLC and UNITED BISCUITS are positioned at similar distances from the average: UNITED BISCUITS sold 1,832,400, about 22 times more than the average, and G. F. LOVELL sold 3,722, about 22 times less. For both, the relative departure from the expected is $e = 1.35 = \log 22$. Only the signs are different.

Financial ratios y/x can be written as a difference in log space:

$$\log y_j - \log x_j = (\mu_y - \mu_x) + (e^y - e^x)_j \quad \text{or, in ordinary space,} \quad \frac{y_j}{x_j} = R \times f_j \tag{3}$$

with $R = \exp(\mu_y - \mu_x)$ and $f_j = \exp(e^y - e^x)_j$. R is the ratio standard: It is the number of times y is expected to be larger or smaller than x. An estimated R is $\exp(\overline{\log y} - \overline{\log x})$, the median of the ratio. f_j is the deviation from the ratio standard: It is the number of times the ratio, observed in firm j, is larger or smaller than the expected for the industry.

This manipulation shows that financial ratios only allow the strict, two-parametric, lognormality of their components. Equation (2) is restrictive in that it doesn't model all the characteristics of raw numbers described in previous chapters. As recalled, the literature also considers ratios as restrictive since they can't model non-linearity and non-proportionality. Empirically, it is possible to write a model incorporating the features unexplained by (2): Three-parametric lognormality (thresholds) and an effect common to all items from the same report. In that case, an observation x_j would now be explained as

$$\log(x_j - \delta_x) = \mu_x + \sigma_j + \varepsilon_j^x. \tag{4}$$

δ_x is the threshold of x, σ_j is the effect of size in report j, and ε^x is the deviation from this effect observed in x. This chapter shows that there is a simple point of view able to explain both (2) and (4), and hence the findings of previous chapters. According to it, there is no contradiction between proportionality as a statistical effect, and non-proportionality in ratios. It also shows how to extend ratios so as to overcome their limitations.

4.1 Ratios and the Common Effect

The Gaussian distribution is often seen as the result of many independent elementary perturbations. This approximation entails the assumption of a constant effect. For example, the probability of getting odds, when tossing a coin, is a constant value of 1/2 no matter the number of games or the size of the coin. And the probability of getting particular proportions of odds when tossing a coin in several sequences of games draws a Gaussian distribution. This constant probability of 1/2 governing the game referred to is what we call a constant effect. If, however, any random change dx suffered by a variable x is proportional to the value of x itself, the effect is no longer constant. It is a proportionate one.

Gaussian variables spread their final realizations in the neighbourhood of an expected value. It's unlikely to find cases many orders of magnitude larger or smaller than the expected. This is because the random changes leading to them are similar. When the random changes leading to any final realization are similar only when taken as proportions of the momentary value of the variable, the distribution of those final realizations exhibits strong positive skewness.

The lognormal probability distribution is consistent with proportionate mechanisms. This fact is known as the Gibrat law [17] (1931). Let x be the position of a stochastic variable. If dx, the random changes affecting x, are expected to be proportional to x itself,

$$\text{the quotient} \quad \frac{dx}{x} \quad \text{is expected to be independent of } x.$$

So, if we can find a function

$$z = f(x) \qquad \text{such that} \qquad dz = \frac{dx}{x} \tag{5}$$

then the new variable z will obey the assumption of a constant effect. In the case of dz being many, independent, perturbations, $f(x)$ is the logarithmic function. Aitchison and Brown [1] (1957) contain a detailed explanation of this reasoning. Notice that lognormality emerges as a result of the Central Limit theorem. The normality of the process governing dz is not required as an assumption, whenever the dz are many, independent, changes.

Since the elementary perturbations dz produce small changes dx in x, expected to be a proportion of x itself, dz can be seen as an elementary relative growth, and z as an expected relative growth. Gaussian final realizations $z_j = \log x_j$ are thus explained by a central trend, μ_x, resulting from a constant effect determining the average relative growth, and by each particular departure from μ_x, the e_j^x, affecting only firm j. The meaning of e_j^x is the same as in (2). Hence, the proportionate effect could be used to explain the lognormality observed in raw data, and has been quoted by McLeay in this context. He suggested a distinction between items reflecting size and the ones which couldn't "be treated as size measures" [23].

Our approach is different. In our study no attempt is made to specify the behaviour of any particular item. We assume that, in the case of raw numbers, the Gaussian relative growth dx/x is the sum of two components: A common and strong one, σ_j, which accounts for random changes acting over report j as a whole, and a weak component, particular to item x. Then, if x and y are the positions of two items from the j^{th} report, dx_σ and dy_σ are random changes in x and y caused by σ_j, a disturbance influencing both. A unique, proportionate, effect would mean:

$$\frac{dy_\sigma}{y} = \frac{dx_\sigma}{x}.$$

When considering the whole sample of $1, \cdots, j, \cdots, N$ firms and many, independent, perturbations leading to final realizations of x and y, this mechanism yields

$$\log y_j - \log x_j = (\mu_y - \mu_x) + (\varepsilon^y - \varepsilon^x)_j \quad \text{or, in ordinary space,} \quad \frac{y_j}{x_j} = R \times f_j \tag{6}$$

ε_j^x and ε_j^y have the same meaning as in (4): They measure deviations from the expected, given the size of firm j. They represent the variability unexplained by the common effect, σ_j, in each component of the ratio. The ratio standard can be estimated as $R = \exp(\mu_y - \mu_x)$, and the multiplicative deviations from it are $f_j = \exp(\varepsilon^y - \varepsilon^x)_j$.

The form obtained in (6) seems similar to equation (3), the one based on empirical manipulation, but it isn't. The residual difference is the same, that is, $\varepsilon^y - \varepsilon^x = e^y - e^x$, but in (3) there was no room for considering the size of the firm. The ε^x are very important in this study. Their ability to measure deviations from the expected for size makes sense on grounds of financial statement analysis.

Notice that, unless we can isolate σ_j, the strong, common, effect, we cannot determine each ε^x and ε^y separately. Conversely, it is impossible to know the value that σ assumes for firm j unless we know ε^x or ε^y. Since both the common effect and the particular ones are hidden from direct observation, the above developments seem useless. However, it is possible to devise a method for overcoming such a limitation. We shall return to this important topic in section 6.1.

Notation: Differences between expected values are written as $\mu_{y/x} = \mu_y - \mu_x$ or, for the ratio standard, as $R_{y/x}$. For example, we often write $\varepsilon^y - \varepsilon^x$ as $\varepsilon^{y/x}$. In the case of deviations of log items from the expected, the e^x, or contrasts between these deviations, the $e^{y/x}$, or deviations from the expected for size, the ε^x, we use superscripts to avoid too many subscripts.

Ratios with more than two items: For more than two items an obvious extension emerges. Let $x_1, \cdots, x_i, \cdots, x_M$ be the position of M items for report j. dx_i are random changes in x_i caused by σ, a common source of variability. Considering the way such disturbance affects the relative growth which is about to generate the set of x_i we can say that

$$\frac{dx_1}{x_1} = \frac{dx_2}{x_2} = \cdots = \frac{dx_M}{x_M} \tag{7}$$

For example, two groups of items $y_1, \cdots, y_k, \cdots, y_K$ and $x_1, \cdots, x_l, \cdots, x_L$ lead to

$$\left[\frac{1}{K}\sum_{k=1}^{K}\log(y_k) - \frac{1}{L}\sum_{l=1}^{L}\log(x_l)\right]_j = \frac{1}{K}\sum_{k=1}^{K}\mu_k - \frac{1}{L}\sum_{l=1}^{L}\mu_l + \left[\frac{1}{K}\sum_{k=1}^{K}\varepsilon^k - \frac{1}{L}\sum_{l=1}^{L}\varepsilon^l\right]_j$$

Despite its outlook, this model is very simple and can be seen just as an expansion of equation (3). Instead of log items, expected values and residuals, we now have averages of log items inside report j. It leads to ratios like these ones:

$$\frac{\sqrt{CA \times CL}}{TA} \quad \text{or} \quad \frac{S}{\sqrt{N \times W}} \quad \text{or} \quad \frac{A \times B \times C}{D^3}$$

Notice that such ratios only require to be statistically valid that the residuals f_j are multiplicative. We must point out that, conversely, additive residuals in ratios, as accepted in the literature and practice, aren't statistically consistent. If residuals were additive their variability wouldn't retain its statistical nature after expansions such those displayed above. This remark applies to the "Pyramid of Ratios" and other well known expansions.

Degrees of freedom involved: The ratios discussed so far engage one free parameter. For modelling expected values one free parameter is enough. But the variability of ratios remains unaccounted for. This has important implications for the interpretation of departures from ratio standards. Ratio standards are insensitive to the distribution of ratios. No consideration of their variability is required. Conversely, disturbances in this variability don't affect ratio standards. The inclusion of more than one item in the ratio components don't account for more explained variability. The number of degrees of freedom engaged remains equal to one. However, more variables, if conveniently selected, can enhance the ability of ratios to recognise features by smoothing out or reinforce some components of the variability of their components.

4.2 Free-Slope Ratios for Prediction

Equation (6) defines ratios as "errors-in-variables" or functional relations. Ratios yield a contrast between two items, both affected by errors. Such a contrast measures how big are the discrepancies between their components. Thus, the above description is intended to assess deviations from standards, not for prediction. This is because ratios don't model the variability of their components.

Ratios intended to predict items must be able to account for the variability of deviations from the standard. In order to do it, more than one parameter is required. The simplest approach to achieve this, consists of introducing in the mechanism leading to ratios a slope, b_i, individualizing the relative growth of each item in a report:

$$b_1 \times \frac{dx_1}{x_1} = \quad b_2 \times \frac{dx_2}{x_2} = \quad \cdots \quad b_M \times \frac{dx_M}{x_M} \tag{8}$$

The b_i are gain or attenuation factors expressing different degrees of linear correlation between the mechanism leading to items. Notice that only $M-1$ of these b_i are independent. In the simple case of b being similar across firms the consideration of two items, y and x, yields free-slope ratios of the form

$$\log y_j - b \times \log x_j = w_0 + (\varepsilon^y - \varepsilon^x)_j \quad \text{or similar. In ordinary space,} \quad \frac{y_j}{x_j^b} = \exp(w_0) \times f_j$$

b and w_0 are parameters. w_0 isn't independent of b.

The multivariate descriptor of this class has the form

$$\sum_{i=1}^{M} w_i \times \log x_i = w_0 \tag{9}$$

in which the residual is omitted. w_i are parameters. (9) is a linear relation. In ordinary space, free-slope ratios yield non-linear, though proportional, relations. This non-linearity mainly affects large firms.

Predicting items with free-slope ratios: Functional relations describe mechanisms. Mechanisms should be plausible. Free slopes in log space aren't plausible as they imply a unique average relative growth for the same item across many firms. Moreover, it would be inadequate to consider non-linearity as a rule. However, when the goal is just prediction, there are no objections to regressing in log space, provided a proper interpretation is given to the obtained model. Figure 13 shows a free-slope ratio predicting Sales in terms of Net Worth.

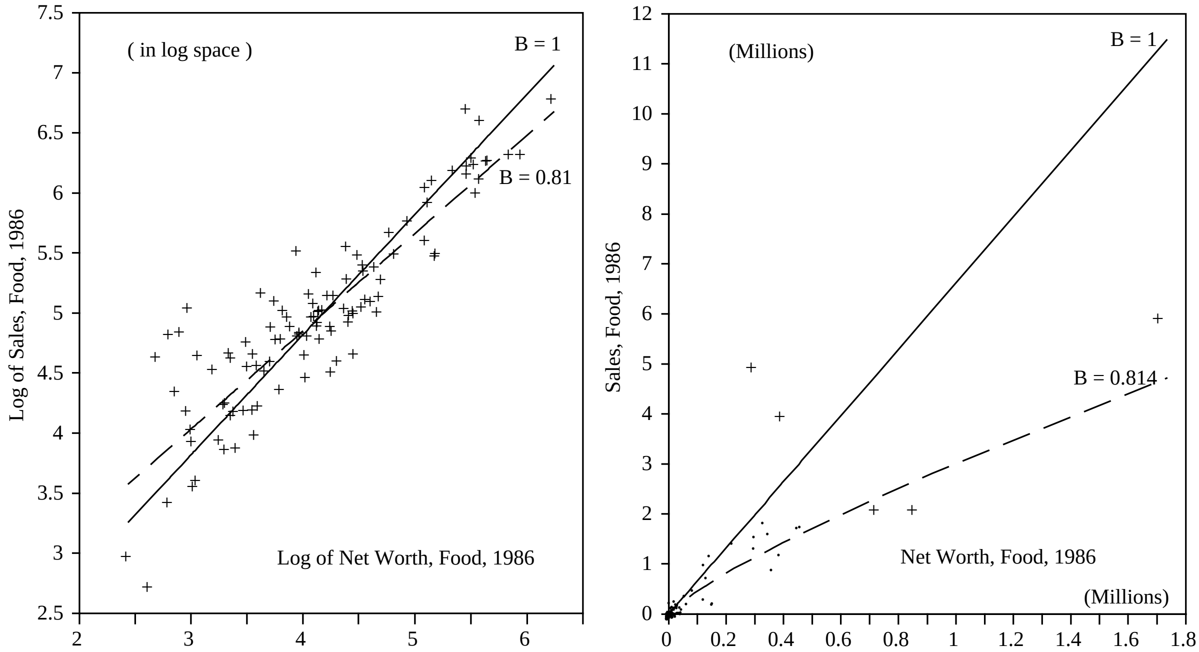

Figure 13: Comparing a free-slope ratio ($B = 0.81$) with the usual one ($B = 1$).

4.3 Non-Proportional Ratios

An usual topic in accounting literature is to call the attention for the assumption of strict proportionality underlying the use of ratios. Such a statement is descriptive. Now we enumerate the assumptions attached to ratios in a generative, rather than in a descriptive way:

1. Items are final realizations of elementary random changes. Such changes, when expressed as proportions of the item they affect, are, in average, the same. This is the Gibrat law.

2. The elementary random changes leading to final realizations of accounting items are, when expressed as proportions of the item they affect, a sum of two components: One, affects in the same way all the items in the same report. Another one, is particular to each item.

The advantage of using a generative description is that we can develop models other than ratios, consistent with this basis. Ratios can now be seen as models that obey the statistical or expected proportionality of random effects, instead of a strict proportionality. Which models are also allowed by such a description? In this section we extend ratios so as to model non-proportionality.

Thresholds and the Gibrat law: The relation $dx/x = dz$ is simplistic as it accepts growing from naught. The Gibrat law points out a more realistic basis by admitting that the random changes dx affecting x are proportional, not to x itself, but to $x - \delta$. We call this δ a threshold. Since the process leading to a realization of x starts with a non-zero value for $x = 0$, the increments x receives on the beginning of its growth are, in average, proportional to such threshold. Therefore,

$$\text{instead of} \qquad dz = \frac{dx}{x} \qquad \text{we should write} \qquad dz = \frac{dx}{x - \delta}$$

to describe the generation of a particular item x. Such a process leads to a class of ratios which can have different characteristics according to the magnitude, sign and position of the thresholds of their components. In some cases, but not in all, these threshold ratios are non-proportional.

Notice that δ shouldn't be taken as the value of x at the beginning of the process leading to its final realization. So long as the mechanism is proportional, initial values don't induce non-proportionality in cross-sections. Non-proportionality emerges only when the random changes dx are proportional to values which are not x.

Next we briefly describe some of the possible models resulting from thresholds.

An overall threshold in the denominator: In the simplest case, δ is a constant affecting all firms in the sample. When the threshold acts on the denominator but not in the numerator, the bivariate version would be

$$\log(y_j) - \log(x_j - \delta) = \mu_{y/x} + \varepsilon_j^{y/x} \qquad \text{In ratio form,} \qquad \frac{y_j}{x_j - \delta} = R \times f_j$$

In the above expression, and in all subsequent ones, the item affected by the threshold — in this case it is x — receives a transformation similar to the one used to achieve three-parametric lognormality. In fact, non-proportional ratios occur when any of the ratio components is three-parametric lognormal.

The non-proportionality introduced in ratios by thresholds is different from the one introduced by a constant term in regressions:

$$y_j = -\delta \times R \times f_j + x_j \times R \times f_j$$

Here, displacements are proportional to residuals. Hence, deviations from proportionality vary from case to case. They are also small, provided δ remains small. The non-proportional term is significant only for values of x_j near δ, that is, whenever the final realizations of items are near the threshold. This could happen when the growth leading to an item is weak (a very small relative growth or very few random changes) when compared with the threshold. Cases far away from δ exhibit proportionality since $x_j \gg \delta \times R \times f_j$.

The examination of bivariate scatterplots of items in log space can detect departures from strict proportionality when they are significant. As figure 14 depicts, the log and the ratio transformations produce a trade-off between non-proportionality and non-linearity, so that departures from proportionality result in a visible bending of the scatter. The small non-linearity observed in figure 5 (left) on page 21 stems from a significant threshold.

An overall threshold in the numerator: By considering a threshold affecting the numerator of the ratio, we get non-proportional terms which can more easily be significant. The expression

$$\log(y_j - \delta) - \log(x_j) = \mu_{y/x} + \varepsilon_j^{y/x} \qquad \text{means the ratio} \qquad \frac{y_j - \delta}{x_j} = R \times f_j$$

which could be written as $y_j = x_j \times R \times f_j + \delta$. This threshold acts as an intercept in a regression. It introduces a constant displacement affecting all cases in the sample. Notice that this model is still not a regression. The difference, however, isn't functional. It stems from the multiplicative nature of the residuals.

Thresholds both in the numerator and in the denominator: When considering δ_y and δ_x as both significant, the amount of non-proportionality in ratios results from their interaction. A reinforcement will occur when δ_y and δ_x have different signs. The overall effect depends on R, the expected proportion. In a particular case, $\delta_y = \delta_x \times R$, both thresholds cancel out. The remaining non-proportionality is case-dependent.

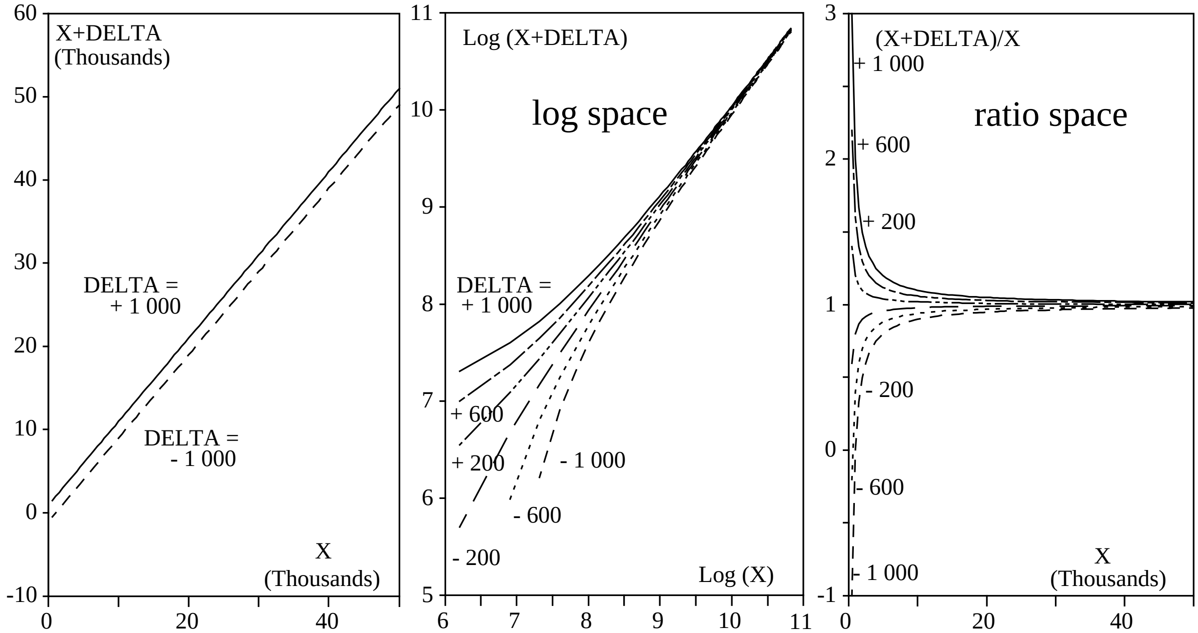

Figure 14: When $Y = X + \delta$ is transformed, the fact that $\delta \neq 0$ introduces non-linearity in the resulting relation. Such non-linearity affects mainly values near δ. On the left, several $Y = X + \delta$ with small δ. In the centre, the same in log space. On the right, in ratio space.

Proportional thresholds: The mechanism leading to the above descriptors requires an overall displacement — a threshold acting upon the whole of the sample. Overall thresholds suppose the existence of overall costs or income and the corresponding cause must be external to firms. We now consider the case of thresholds which are internally-generated. Mechanisms internal to the firm are likely to generate thresholds proportional to size. For $1, \cdots, j, \cdots, M$ firms, δ_j is now a particular threshold concerning the proportionate effect leading to each x_j. This threshold will act as a new variable, not as a parameter. Hence, the model collapses into the no-threshold ones. In fact, if δ_j is proportional to the size of the firm, it is similar to any other accounting item. For instance, we could write

$$\delta_j = x_j \times R_{\delta_j} \times f_{\delta_j} \qquad \text{and we would have a relative growth} \qquad \frac{dx}{x \times (R_{\delta_j} \times f_{\delta_j} + 1)} = dz$$

for the generating process of a particular realization of x. Since R_{δ_j} and f_{δ_j} are not involved in the growth of x, the resulting model is a version of the free-slope ratio. thresholds proportional to the size of the firm don't break proportionality. They just induce differences in the way each item is affected by the common effect.

The described model is interesting because thresholds internally-generated have been often used in the literature as an example of the plausibility of intercept terms in bivariate relations. It was an awkward choice since, as we see, thresholds acting just as another item aren't likely to induce overall translations. We now analyze this subject using a different point of view.

4.4 The Basis for the Existence of Non-Proportionality

Thresholds can have causes internal or external to the firm. The most general one is internal: when nonexistent variables ought to exist. All growth starting with $x = 0$ must have at its origin a threshold acting like a seed since naught can't grow. Notice also that, owing to its exponential nature, the final realizations of proportionate growth are likely to attain values many orders of magnitude larger than this seed. In such cases, $x - \delta \approx x$ and the non-proportional term vanishes.

Internally-generated thresholds: The foundation invoked in some texts to justify the existence of significant departures from proportionality is coincident with the model we call the proportional threshold. For example, Lev and Sunder [22] (1979) say that

> The relationship between gross profit and sales probably contains a positive constant term given the frequent existence of a significant fixed costs component. Accordingly, observed differences in gross margin ratios (over time *or across firms*) reflect the confounding effects of differences in efficiency, differences in the level of fixed costs, and sales.

This literature says that, because in each individual firm some internal mechanisms exhibit constant terms, the corresponding statistical variables, obtained when gathering many firms in a sample, would exhibit also a constant term. But, as we saw above, this is not so. This pitfall seems to be another case of picturing time-series while working with cross-sections. The meaning of a cost being fixed is that it is fixed inside a firm. But it can be variable across firms. As a first approximation, large firms exhibit large fixed costs, small firms exhibit small fixed costs and infinitesimal firms would exhibit infinitesimal fixed costs. On the limit, the zero-sized firm should have zero fixed costs thus yielding strict proportionality. Whittington [30] (1980) clearly distinguishes between time-series and cross-sections when addressing this problem:

> In cross-section, such an interpretation (sales-unrelated income) could not be placed on the constant term: It would now represent an estimate of the average amount of sales-unrelated income for the average firm, provided the further assumption is made that "sales-unrelated income" is strictly independent of size.

This statement is equivalent to ours. In cross-section, Fixed Costs should be regarded as another item with nothing special about it.

Overall thresholds: Mechanisms internal to the firm don't generate thresholds. Only the ones acting over all the population can do it. But can they induce strong thresholds? It seems as if there is a limit for the plausibility of displacements affecting entire samples. Clearly, if an overall cost were big enough to be noticed by large firms, it would be far greater than the earnings of the smaller ones, leading them to insolvency. And if it were small enough to allow any firm to survive, it would be unnoticed by most of them and its effect would be negligible. For example, a fixed cost of 3,722 thousand pounds over the whole of the Food Manufacturers industry in the UK, would represent to UNITED BISCUITS just 0.2% less earnings in 1987. But the same cost would eat up the sales in G. F. LOVELL PLC. All the firms similar or smaller in size would perish (about 5% of the industry).

Summary: In this section we have shown that non-proportionality in ratios cannot have its origin in mechanisms internal to the firm, like fixed costs or income. Also, given the lognormality of the ratio components, this non-proportionality should be small.

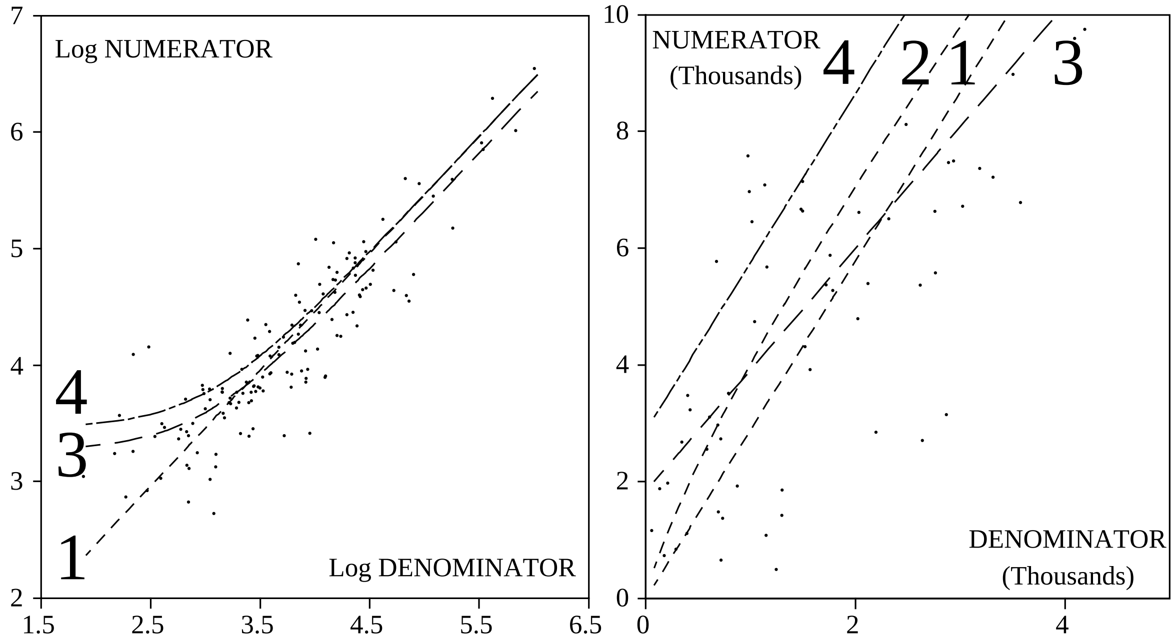

Figure 15: The typical shape of free-slope and threshold ratios. On the left, log space. On the right, a magnification of the region near the origin in ordinary space. (1) is the usual ratio, (2) is a free-slope one, (3) and (4) are threshold ratios. Electronics, 1987, FA/CA.

4.5 Using Extended Ratios

This section shows an example of the applicability of the extended ratios. We selected Electronics, 1987, and the ratio FA/CA. This is because in some industries CA is three-parametric. The choice of FA for the numerator is explained in section 5.5.

Five different models are compared: The numbering (1 to 5) refers to figure 15 on page 46. Next we describe each model. As usual, x represents the denominator and y the numerator of the ratio.

Model 1: The usual ratio. It engages one degree of freedom. Its unique parameter, the median in log space, is estimated as a by the Least-Squares model

$$\log y_j = a + \log x_j + \varepsilon_j^{y/x} \qquad \text{yielding the ratio} \qquad \frac{y_j}{x_j} = 10^a \times f_j$$

The graphical representation has the label 1 in figure 15. It is a 45^o straight line in log space. In ordinary space it is also a straight line passing through the origin.

Model 2: The free-slope ratio. It engages two degrees of freedom. Its two parameters are a and b, the slope. They are estimated by the Least-Squares model

$$\log y_j = a + b \times \log x_j + \varepsilon_j^{y/x} \qquad \text{yielding the ratio} \qquad \frac{y_j}{x_j^b} = 10^a \times f_j$$

The graphical representation has the label 2 in figure 15 (on the right). This model has already been displayed in figure 13 on page 40. It is a straight line in log space. It is non-linear in ordinary space. It goes through the origin.

Models 3 and 4: The threshold ratios. They engage two degrees of freedom. But in model 3, both parameters (a and δ) are estimated jointly. In model 4, a is taken as known. Then δ is estimated

Model	a	b	δ	R^2	Skewness	Kurtosis
1	-0.46			72%	0.27	2.18
2	-0.21	0.94		76%	0.07	1.30
3	-0.49		528	78%	-0.12	1.19
4	-0.46		327	76%	-0.12	1.78
5	-0.40	0.99	403	78%	-0.12	1.19

Table 13: The estimated parameters and statistics for five ratios: (1) is the usual ratio, (2) is a free-slope one, (3) and (4) are threshold ratios, and (5) is the threshold plus free-slope ratio. Electronics, 1987, FA/CA.

based on this assumption. They are estimated by the Least-Squares model

$$\log y_j = a + \log(x_j - \delta) + \varepsilon_j^{y/x} \quad \text{yielding the ratio} \quad \frac{y_j}{x_j - \delta} = 10^a \times f_j$$

The graphical representations have the labels 3 and 4 in figure 15. They are parallel and non-linear in log space and straight lines in ordinary space. They don't go through the origin. Model 4 converges to the ordinary ratio for medium-sized and large firms.

Model 5: The threshold plus free-slope ratio. This ratio is omitted in figure 15. It engages three degrees of freedom. The parameters are a, b and δ. It is the result of considering free slopes and thresholds together. The parameters are estimated by the Least-Squares model

$$\log y_j = a + b \times \log(x_j - \delta) + \varepsilon_j^{y/x} \quad \text{yielding the ratio} \quad \frac{y_j}{(x_j - \delta)^b} = 10^a \times f_j$$

Next table shows the obtained parameters along with the variability explained, the skewness and the kurtosis of the residuals in log space (the $\varepsilon_j^{y/x}$).

Model 1, the simple financial ratio, accounts for 72% of the variability. By letting the slope vary freely, further 5% is explained. Both the skewness and the kurtosis of residuals improve. In the 3^{rd} and the 4^{th} models, the gain in explained variability is similar to the one obtained with free slopes. The explained variability didn't improve in the 5^{th} model.

The threshold ratio explains as much variability as the free-slope plus threshold model (the 5^{th}). It is interesting to notice that the 2^{nd} model uses the free-slope to approach the effect of the threshold. Once such threshold is accounted for, the slope returns to the value of 1 (5^{th} model). The threshold is itself very small. When using the method described in section 1.1 to estimate the thresholds in CA and FA, we obtained values which agree with the above ones. CA is three-parametric lognormal. Significant departures from normality vanish for $\delta > 300$. The maximum W is obtained with $\delta = 570$. FA is two-parametric. The kurtosis of residuals is strong and it will not vanish with thresholds or free slopes.

From the four extensions of ratios, the most attractive one for financial analysis seems to be the 4^{th}. It accounts for non-proportionality but it approaches the usual ratios for larger values of its components. It is simple to implement and robust regarding influential cases.

4.6 Discussion and Conclusions

In this chapter we discussed the validity of ratios in the light of previous findings. We showed that there is a stochastic mechanism able to integrate such findings in the same overall explanation. According to it, ratios can be extended in several ways. Firstly, they can have more than two components. The sole requirement for the statistical validity of such ratios is the use of multiplicative residuals. Ratios can also be viewed in log space as a regression. Such free-slope ratios preserve proportionality. They introduce non-linearity for large firms. Finally, the existence of thresholds eventually introduces non-proportional relations between components of ratios.

Distortions in proportionality resulting from overall thresholds depend on several factors. They are maximal for thresholds in the numerator of the ratio or when the signs of the thresholds of the numerator and the denominator are different.

Threshold ratios seem promising for financial statement analysis and statistical modelling. They are robust, easy to estimate, and it is likely that they will be able to gather in one simple model the correct relation between two items for firms of very different sizes. Free-slope ratios are adequate for prediction. They shouldn't be used to assess deviations from standards.

Correctly estimated expected values of ratios are insensitive to their distributions. The free-slope and, to a smaller extent, the threshold ratios, suffer misleading influences in the presence of constraints.

The common effect: According to our approach, all the items of the same report should be expressed in terms of a common effect, size, and deviations from it. A common effect greatly simplifies the formal treatment of modelling with lognormal data. We have shown that non-proportionality is compatible with this assumption. In section 6.1 we shall see that a simple manipulation of equation (4) allows size and deviations from it to become available as statistical variables.

The extension of the effect of size to items like Working Capital, Earnings and Funds Flow isn't usual in the literature. The common effect itself, is also a new way of interpreting size. But our findings dismiss any strong differentiation between items which are accumulations and those which aren't. The cross-sectional lognormality observed in items is determined by differences in size, not by mechanisms specific to a few items. We don't think Fixed Assets is more lognormal than Earnings owing to the proximity of the former to the assumptions of the Gibrat law. As stressed above, we picture proportionality as a stochastic effect explaining differences in size observable in cross-sections containing many firms.

The Gibrat law explains lognormality in raw data and the existence of non-proportionality in ratios. It is also plausible as a mechanism of accumulation and it ensures a consistent development of extensions of ratios. We used it here as a point of view able to gather the main facts about ratios in a unique formulation.

Chapter 5

The Distribution of Ratios

The deviations from ratio standards, the $f^{y/x}$ in formula (6), should be positively skewed due to their multiplicative nature. The same should be true in the case of the ratio output, $R \times f^{y/x}$. In practice, some ratios are Gaussian or even negatively skewed. How is it possible? This chapter applies the findings of the previous ones to the problem of the distribution of ratios.

We introduced the literature on this subject early in this study. The problem of the distribution of ratios concerns a vast amount of different situations. There are many possible ratios, many possible choices for the definition of samples, many tests and criteria to analyze the results. The studies on the distribution of ratios typically try to avoid dispersion by using Deakin's set of 11 ratios. Despite this effort the results are difficult to interpret. Our method consists of noticing that ratios are bivariate relations. In section 5.1 we examine the influence of external forces on the bivariate lognormality of raw data. This allows us to predict which ratios are near normality, which ones have negative skewness and which ones remain broadly lognormal. Our predictions are supported by the published research. In section 5.2 we describe regularities observed in the behaviour of ratios. As a result, we suggest procedures for the selection of variables in statistical models.

Since this study is also intended to cover the statistical modelling of accounting relations, we selected items according to criteria which somehow differ from those adapted in the research concerned with the distribution of ratios. However, we didn't go too far in the differentiation from the published studies. It seems desirable to compare our results with other's. For example, we used the ratio output, $R \times f^{y/x}$, instead of the deviations from ratio standards, the $f^{y/x}$. Also, some of the selected ratios are usual in the literature.

5.1 The Effect of External Constraints

Usually ratios exhibit strong positive skewness. This is consistent with their multiplicative nature. However, the literature mentions ratios which are Gaussian or even negatively skewed. Typically, TD/TA is reported as being Gaussian (see Deakin [10], 1976, Ezzammel and Mar-Molinero [13], 1990, p. 11). The reason for this is straightforward. Accounting identities like $TA = CA + FA$, make it impossible for some bivariate relations to have all the values a skewed distribution would allow. Such identities act as a constraint introduced in the normal course of their variability. This effect, mentioned in the literature to explain why some ratios are bounded, had never been associated with the strong departures from positive skewness observed in the distribution of ratios. In log space it turns out that

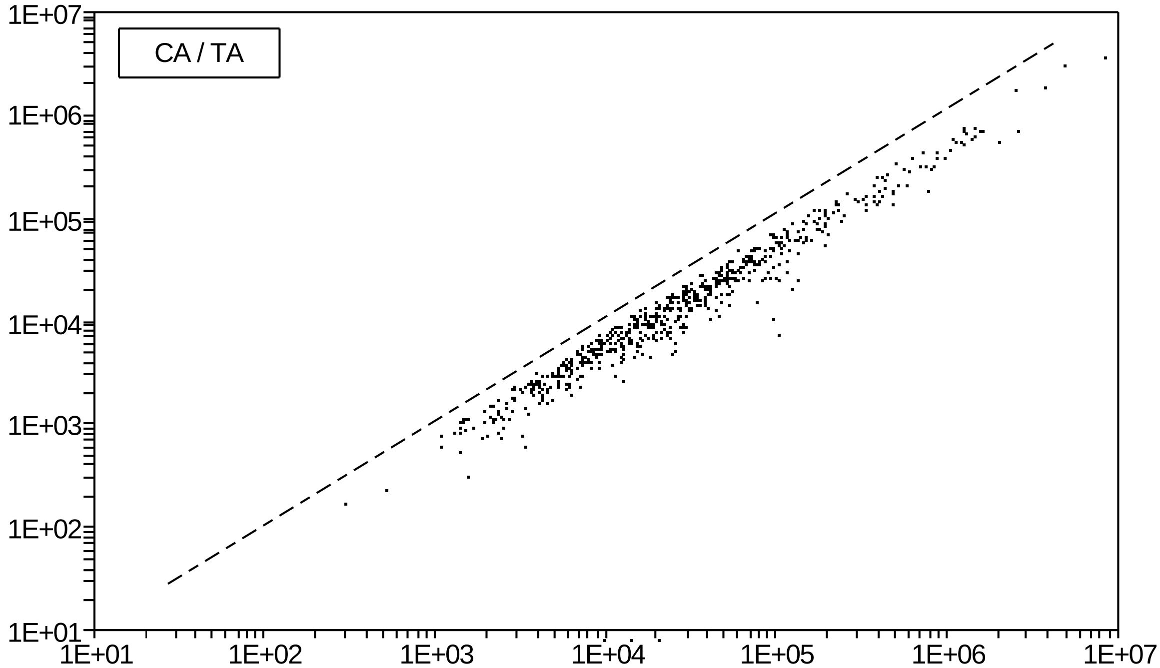

Figure 16: Scatterplot in log space showing the effect of a strong constraint imposed on CA (Y-axis) by TA (X-axis). All groups together, 1984. The dashed line is CA = TA.

this effect is clearly observable and self-explanatory (see figure 16).

We say that there is a constraint if, due to any accounting identity or other external force, the bivariate relation $y_j/x_j = R \times f_j$ is bounded so that one of the next non-equalities hold.

$$\text{for any } j, \quad x_j > y_j \quad \text{or} \quad y_j > x_j$$

The non-equality on the left can be found in constrained ratios where the numerator is bounded by the denominator. An example is Debt to Total Assets. The non-equality on the right arises in ratios in which the denominator is bounded by the numerator. This isn't as frequent as the first case. But, of course, it is possible to create such a situation just by inverting one of the former ratios.

The consequences for the distribution of ratios are different in one case and in the other: When the constraint is $x_j > y_j$, ratios cannot be larger than 1. This frontier is the dashed line, $x = y$, in Figure 16. The effect of this constraint on the distribution of ratios is that of not allowing the spread out of its otherwise positively skewed distribution. Instead of the large, lognormal-like tail, such ratios exhibit a much smaller one. This explains why some studies didn't find positive skewness in a few ratios. We shall see that this constraint can be effective in creating Gaussian-like distributions.

When the constraint is $y_j > x_j$, ratios cannot be smaller than 1. The large, lognormal-like tail towards large values is left untouched but the one towards the small values is now truncated. This increases even more the positive skewness of ratios. According to this mechanism, a strong positive skewness should emerge after inverting one of the apparently Gaussian ratios.

An example: Table 14 on page 51 displays the skewness and kurtosis of three constrained ratios. CA/TA is so strongly constrained that its distribution becomes skewed in the negative direction. Lognormal distributions are two-tailed. If the large tail completely vanishes, the small one introduces

Ratio		Not inverted 1983	1984	1985	1986	1987	Inverted 1983	1984	1985	1986	1987
CA/TA	SKEW	-0.41	-0.43	-0.35	-0.50	-0.59	18.7	5.04	6.29	20.7	17.9
	KURT	0.33	0.26	-0.10	0.33	0.53	377	37.0	71.5	483	337
NW/TA	SKEW	-0.15	0.01	0.01	-0.01	-0.06	17.7	15.4	22.3	12.9	12.4
	KURT	-0.23	-0.14	-0.14	-0.04	-0.04	356	260	536	193	179
FA/TA	SKEW	0.41	0.43	0.35	0.50	0.59	11.2	21.9	23.8	12.7	11.2
	KURT	0.33	0.26	-0.10	0.33	0.53	147	519	592	197	151

Table 14: The skewness and kurtosis of three ratios and their inverse during the period 1983-1987 for all industries together.

negative skewness. NW/TA is almost Gaussian: The large tail doesn't vanish but becomes much shorter so as to balance the small one. Finally, FA/TA remains positively skewed but less than the expected. Its distribution is symmetrical to the one of CA/TA. In the same table we also show the skewness and kurtosis of ratios similar to the above ones but inverted. These values are lognormal-like as the constraint is now working in the same direction as the lognormal skewness.

5.2 Comparing Constrained and Non-Constrained Ratios

The last section was devoted to the identification and description of the effect that external constraints can have in the distribution of ratios. We discussed a limited example, showing how an accounting identity can hide the multiplicative nature of ratios. In this section, apart from providing a more systematic evidence on such effect, we show that ratios are broadly lognormal.

How ratios are affected by constraints: We are interested, not only in describing the constraint mechanism, but also in predicting the skewness of ratios. This is more easily done in log space. When the constraint is $x_j > y_j$, then $\overline{\log y} - \overline{\log x} < 0$, and we must have, for any j, $\epsilon_j^{y/x} < -(\overline{\log y} - \overline{\log x})$. That is, large positive deviations from the expected aren't allowed. When the constraint is $y_j > x_j$, then $\overline{\log y} - \overline{\log x} > 0$, and we must have, for any j, $\epsilon_j^{y/x} > -(\overline{\log y} - \overline{\log x})$. That is, large negative deviations from the expected aren't allowed. Since, in both cases, a constraint prevents $\varepsilon^{y/x}$ from spreading across $\overline{\log y} - \overline{\log x}$, this difference can be used to estimate the extent to which constraints affect the symmetry of the log distribution of ratios. We define

$$\zeta = \frac{\overline{\log y} - \overline{\log x}}{\sqrt{\mathrm{VAR}(\varepsilon^{y/x})}} \tag{10}$$

as the distance, in standard deviation units, separating the constraining frontier from the expected value of the ratio. Thus, for $|\zeta| > 3$, the constraint will be very small. For $3 > |\zeta| > 2$, the constraint will be small. For $2 > |\zeta| > 1$, the constraint will be strong. The severity of a constraint increases with the proximity between the expected values of the ratio components, and it decreases with the standard deviation of the ratio.

Constrained ratios: We examined 14 ratios (see list in table 15 on page 52) formed with items from the Balance Sheet and 2 from the Profit and Loss Account. For such ratios there is an accounting identity or at least a constraint influencing their distribution in a variable extent. Table 15 also lists the skewness and the estimated $|\zeta|$ for each one of those ratios. Apart from the usual abbreviations, Q stands for $CA - I$ and TD for $DEBT + CL$. In ratios built with items having zero or negative-valued

Ratio	1983		1984		1985		1986		1987	
	skew	zeta	skew	zeta	skew	zeta	skew	zeta	skew	zeta
CA/TA (1)	-0.41	1.29	-0.43	1.05	-0.35	1.58	-0.50	1.34	-0.59	1.14
CL/TA (x)	0.636	2.2	0.587	1.84	0.616	2.21	0.56	2.24	0.597	2.26
C/CL (2)	-0.09	1.54	-0.15	2.02	-0.21	1.56	-0.27	1.63	-0.27	1.72
C/TA (3)	1.247	2.88	1.227	2.61	1.357	2.94	1.191	2.87	1.137	2.96
DB/TA (4)	1.873	2.37	1.92	2.29	2.103	2.36	1.774	2.17	2.056	2.21
FA/TA (5)	0.41	1.69	0.434	1.96	0.352	1.76	0.502	1.76	0.592	1.81
I/CA (6)	0.264	1.53	0.102	1.74	0.218	1.57	0.433	1.44	0.294	1.33
I/TA (7)	0.569	2.08	0.473	1.9	0.867	2.16	0.913	2.04	0.812	1.89
NW/TA (8)	-0.14	1.61	0.012	1.85	0.009	1.73	-0.01	1.8	-0.06	1.82
Q/CA (9)	-0.26	1.34	-0.10	1.17	-0.21	1.81	-0.43	1.4	-0.29	1.3
Q/TA (0)	0.504	2.07	0.583	1.92	0.51	2.41	0.387	2.17	0.455	2.26
TD/TA (o)	0.349	2.31	0.173	2.19	0.222	2.23	0.198	2.17	0.177	2.21
EB/S (+)	2.03	3.06	1.63	3.16	2.06	3.12	1.824	3.18	1.480	3.25
W/S (*)	0.42	2.14	0.42	2.01	0.412	2.08	0.395	2.13	0.371	2.13

Table 15: The values of $|\zeta|$ (zeta) and skewness for 14 ratios likely to be constrained in their distributions. $Q = CA - I$; $TD = DEBT + CL$.

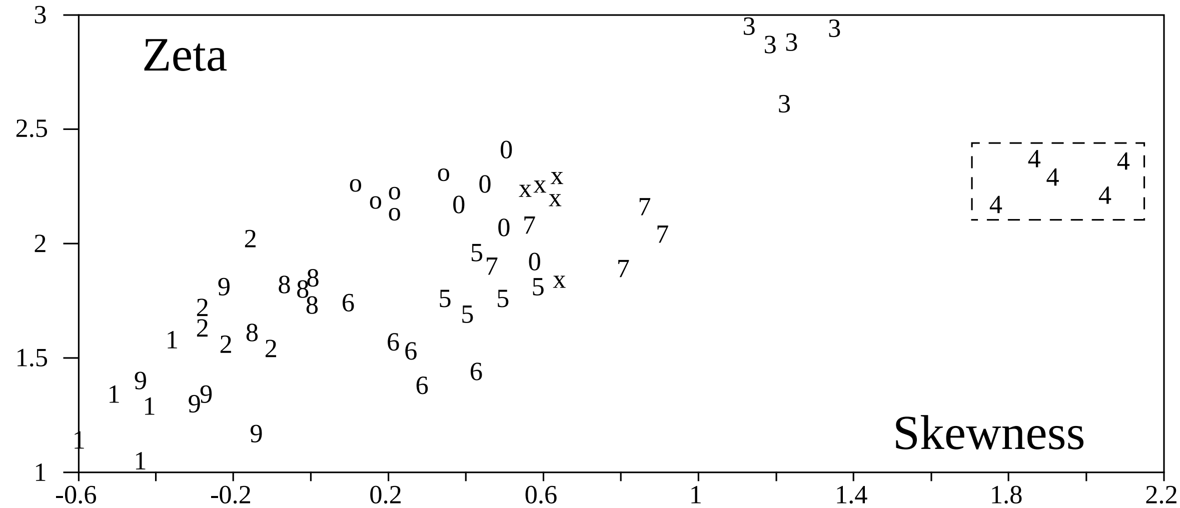

Figure 17: Correlation between skewness and ζ in 14 constrained ratios, 1983-1987.

cases, only the positive ones were used. We gathered in the same sample all the 14 industrial groups. Five years (1983-1987) were checked.

The results show that $|\zeta|$ predicts, to a large extent, the skewness of constrained ratios. For the displayed cases, $|\zeta|$ accounts for 64% of the variability of the skewness ($r = 0.8$). Thus, accounting identities do explain the deviations from a positively skewed distribution in ratios. The imperfection of $|\zeta|$ in approaching the skewness stems from relying on an estimated standard deviation which is itself influenced by the constraint.

In figure 17 (page 54) we reproduce table 15 as a scatterplot. The marks identifying each ratio can be found in table 15, on the left. Figure 17 shows that CA/TA is negatively skewed. I/CA, NW/TA, C/CL, Q/CA and TD/TA aren't far from the Gaussian distribution in what concerns their skewness. Others, like I/TA, approach a skewness of 1. DB/TA, C/TA, have their skewness above 1.

There are other external forces likely to distort the distribution of ratios. Instead of defining frontiers which are impossible to cross, as in the case of accounting identities, these other forces impose frontiers in which only a decrease in the density of cases is observed. For example, the non-equality $CA > CL$ defines one of such decreases because firms avoid, if they can, negative Working Capital.

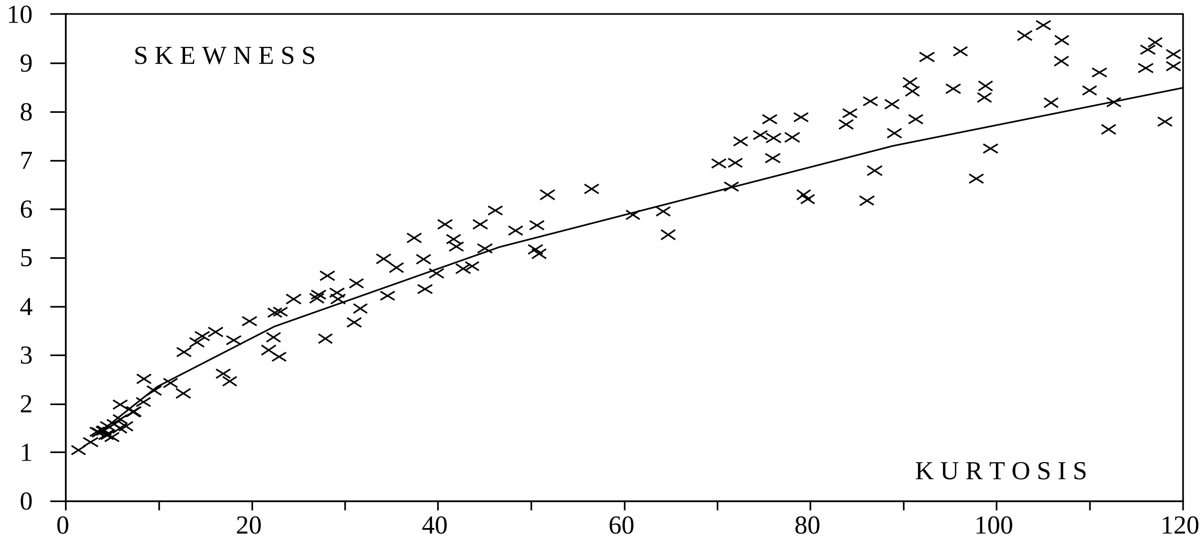

Figure 18: Scatterplot of the skewness with the kurtosis of unconstrained ratios and their inverse. The solid line is the functional relationship linking the skewness of lognormal distributions with the kurtosis.

Non-constrained ratios: The distribution of ratios, when not constrained, has a broad trend towards lognormality owing to its multiplicative nature. Notice that ratios don't have to be lognormal just because their components are. In fact, when examined in log space, ratios show a persistent deviation from the Gaussian distribution, leptokurtosis.

We examined 20 other ratios (and their inverse) for which there is no obvious accounting identity constraining their distribution. These ratios are listed in table 16 on page 56. In both cases ratios were selected so as to represent a random choice amongst all possible combinations. The Number of Employees was often used because there is no obvious constraint affecting its relation with other items. In ratios built with items having zero or negative cases, only the positive ones were used. We gathered in the same sample all the 14 industrial groups listed in table 2, page 14. This sample was then examined for a period of five years (1983-1987).

In lognormal variables the skewness and the kurtosis are not independent. One is a function of the other. Therefore, in order to check the existence of a broad lognormal trend in non-constrained ratios it is enough to plot their skewness against their kurtosis. In case of lognormality, a functional relation should emerge.

Tables 32 and 33 in appendix, contain the skewness and kurtosis of the mentioned unconstrained ratios and their inverses. Figure 18 on page 55 is a graphical reproduction. It displays the regular curve unconstrained ratios and their inverse form when their skewness is plotted against their kurtosis. This regularity obeys the relation between the skewness and the kurtosis of lognormal variables for varying standard deviations. Notice that the SPSS-X package we used in this study, computes the skewness and kurtosis in a way that is not exactly the one found in text books.

In the 200 examined samples (20 ratios and their inverse during 5 years) only three yielded values of the skewness and kurtosis which didn't obey the above relationship. They were from the same ratio, CL/Q, or its inverse, during the years 1983, 1985 and 1987. We further formed a few more ratios with $Q = CA - I$ and we found three other cases of irregular behaviour. They were the ratio $EBIT/Q$ in 1987, and W/Q, in 1986 and 1987.

Ratio	1983		1984		1985		1986		1987	
	skew	kurt	skew	kurt	skew	kurt	skew	kurt	skew	kurt
D/C	-0.7	9.2	-0.8	8.6	-0.1	10.0	0.2	9.3	-0.6	6.2
CA/CL	-1.2	11.0	0.4	2.9	0.8	3.1	0.6	5.9	-0.5	10.5
C/I	0.4	2.1	0.6	2.2	1.1	8.3	0.9	6.7	1.8	11.0
Q/CL	-1.9	12.0	-0.3	4.3	-0.3	3.2	-0.5	5.2	0.1	2.8
W/N	0.3	12.0	-0.1	0.8	-0.4	2.8	-0.3	2.3	-0.6	4.4
S/TA	0.8	4.5	0.8	7.0	1.0	5.4	1.0	5.5	1.2	4.9
S/FA	1.8	5.4	2.0	8.5	1.9	6.9	1.9	5.9	1.6	4.6
S/NW	1.5	4.5	1.6	5.6	1.8	7.8	1.3	4.2	1.6	6.4
S/I	1.3	4.7	1.2	4.6	1.7	8.8	2.3	13.9	3.4	21.0
EB/TA	-2.1	11.2	-1.3	4.9	-1.2	2.9	-1.1	3.1	-1.7	7.9
EB/NW	-1.6	11.4	-0.1	6.1	-0.3	3.6	-0.5	2.7	-0.8	7.3
I/D	-0.6	6.5	-0.2	6.7	-0.3	4.8	-1.6	13.4	-2.0	15.0
W/I	0.1	6.0	0.2	3.2	0.6	3.1	1.2	6.4	1.4	7.8
EB/FA	-0.9	6.0	-0.3	2.8	-0.5	2.8	-0.8	2.2	0.1	3.4
S/N	1.7	5.4	1.8	5.8	1.2	3.3	1.5	4.3	1.3	3.6
EB/N	-0.3	3.9	0.0	1.6	-0.3	2.2	-0.2	1.2	-0.5	2.5
NW/N	0.3	2.2	0.5	1.1	-0.4	3.1	0.0	0.9	0.0	0.9
NW/TA	-1.6	5.6	-1.5	5.6	-0.9	2.2	-0.6	1.5	-0.7	1.5
NW/DB	0.7	1.7	0.8	2.3	0.5	0.5	0.6	0.3	0.8	0.8
DB/S	-0.9	0.9	-1.2	2.4	-0.7	0.7	0.9	1.3	-0.8	0.9

Table 16: The skewness and kurtosis in log space of 20 unconstrained ratios, for a period of five years.

Inverting non-constrained ratios: Constrained ratios change completely their characteristics when inverted and become broadly lognormal. The non-constrained ones remain near lognormality in both situations. They just move along the formal line linking skewness with kurtosis.

5.3 The Log-Leptokurtosis of Ratios

Despite the findings of previous section suggesting a broad lognormality in unconstrained ratios, hardly any of the studied ones is exactly lognormal. When examined closely in log space, they exhibit positive kurtosis of varying severity. This kurtosis was observed in all but one of the studied log ratios. Ratios formed with non-accounting items related to size like the Number of Employees also exhibit it. When sampling by industry, the residual kurtosis will not vanish. As a result, the Shapiro-Wilk test seldom finds a non-significant departure from normality in the log deviations from ratio standards (the $\varepsilon^{y/x}$). We recall from chapter 1 that log items exhibit positive kurtosis as well, but in a smaller degree.

Table 16 contains the usual log statistics for the non-constrained ratios used above. In log space there is no difference in the behaviour of a ratio and its inverse. Distributions are a mirror-image of each other. Therefore, the skewness of the ratios which are the inverse of those displayed in table 16 will simply be the same value with inverted sign. The kurtosis will be the same.

The findings of previous studies: One of the findings of studies in this subject is that the log transformation seems to be unable to improve the normality of ratios. In our opinion this is a result of using precise criteria to assess phenomena which are only broad trends. For example, if we use accurate tests like the Shapiro-Wilk's to measure the lognormality of ratios, we get the general impression that ratios are far away from lognormality. Its precision conceals broad trends.

The use of all sorts of transformations to assess the distribution of ratios only complicates things. For example, if we replicate with ratios the experiment carried out in chapter 1 — which consisted of

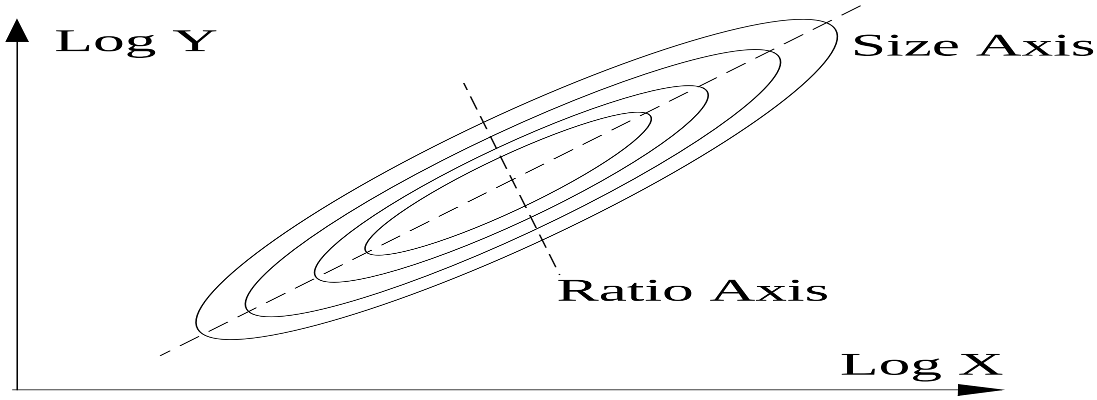

Figure 19: Two-dimensional view of a bivariate density surface in log space.

using progressively higher roots to transform items and observing the results of applying the Shapiro-Wilk or other tests — the results would be confusing. For unconstrained ratios, the skewness would probably diminish with increasing roots but the kurtosis would emerge after some improvements. For some constrained ones the skewness would change sign, becoming negative. Ezzamel, Mar-Molinero, and Beecher [14] (1987) observed this.

5.4 The Two Sources of Statistical Behaviour

In this section we show that there is no contradiction in the fact that the logs of raw data are Gaussian and the logs of ratios are leptokurtic. Let us write equation (6) in this way:

$$\varepsilon_j^{y/x} = (\log y_j - \mu_y) - (\log x_j - \mu_x)$$

That is, in log space, the deviations from the ratio standard, the $\varepsilon^{y/x}$, are a difference of two log items, both with their central trend accounted for. Their distribution is the result of the subtraction of two Gaussian distributions with the same mean. But these two distributions are very similar in spread. A large fraction of the variability of items comes from the strong, common, effect they share. The remaining variability is the source of the positive kurtosis in log space. It is so small a proportion of the total one that it could have any distribution without greatly affecting the overall lognormality of items. In ratios, however, it is prevalent.

We can have a graphical view of this reasoning by considering the bivariate log distribution drawn by the components of ratios. Such distribution is an oblong hill-shaped surface oriented in the 45^o direction and centred in $\mu_y - \mu_x$. The density of cases determines the height of each point in the surface (see figure 19). Such a surface is very thick in one of its main dimensions and very thin in the other one. The largest dimension accounts for most of the variability. In figure 19, the largest dimension is labelled the "Size Axis" and the smallest one the "Ratio Axis". The variability of log ratios is explained by the smallest dimension, the ratio axis. It is orthogonal to the size one, which accounts for the variability introduced by the common effect.

When an observer positions himself so that the largest dimension of this surface becomes parallel to his horizon, he sees a Gaussian shape. When he observes it transversally, it yields a leptokurtic shape. Thus, the weak, particular, effect is the source of leptokurtosis in accounting data, and the strong, common, one is the source of their lognormality. For example, the small amount of kurtosis observed in the logs of raw numbers denotes the influence of their own variability superimposed to a common, Gaussian, one. In section 6.1 we shall return to this topic.

5.5 Avoiding Asymmetry in Ratios

This section suggests a way of avoiding the asymmetry introduced in ratios by external constraints.

The proportionality of ratios is understood as a statistical quality related to the non-existence of significant constant terms in cross-sectional relations between the numerator and the denominator. Here we recall a different meaning, concerning the formal relation between numerator and denominator, not the statistical one. A quotient is said to be a proportion when the numerator is part of the denominator. Relative frequencies or probabilities are proportions. All the ratios bounded by the denominator are proportions in this sense. Therefore, it seems wise to apply to such ratios the well-known recipes to deal with similar cases. The simplest of such recipes is the "odds-ratio":

$$\text{For proportions } p_i = \frac{x_i}{\sum_i x_i} \text{ there is an "odds ratio" defined as } o_i = \frac{p_i}{1-p_i} = \frac{x_i}{\sum_i x_i - x_i}.$$

When the numerator of a ratio is bounded by the denominator, this transformation incorporates the underlying accounting identity, thus yielding a new, unbounded, variable. The odds ratio corresponding to FA/TA is the ratio FA/CA. The one of NW/TA is $NW/(DEBT + CL)$. The difference between odds-like ratios and the corresponding proportion-like ones is just functional. The information contained in both is the same. Thus, it is possible to avoid ratios affected by constraints by using the corresponding odds ratio instead. This solution only applies to ratios constrained by an accounting identity. In the Profit and Loss account, Sales is not a total. There are other sources of income. But, for industrial firms, it acts almost as if it were. Instead of OPP/S we can use $OPP/COGS$. This new odds-like ratio is unconstrained.

5.6 Summary

In this chapter we studied the distribution of ratios. We found a broad trend towards lognormality, as expected. However, a few factors affect the final distribution that particular ratios assume. Firstly, accounting identities and other external forces can act as constraints, hiding their multiplicative nature. This factor induces the severe deviations from lognormality reported in the literature for ratios like NW/TA and TD/TA. Apart from accounting identities, ratios are also affected by managerial practice and by other external forces.

Secondly, when observing in log space residuals which are broadly lognormal, a persistent leptokurtosis emerges. The weak, particular, effect is the source of this log positive kurtosis. The strong, common, one, can be identified as the source of the Gaussian behaviour of accounting data.

Chapter 6

Size And Grouping

Size and grouping are the main sources of variability in financial statement data. In this chapter we study both. Firstly we discuss the building of an estimated common effect for assessing the size of the firm. Secondly we suggest two techniques to measure the homogeneity of groupings and the complexity of the randomness present in accounting data because of groupings.

6.1 Estimating Size: The Case-Average Model

This section shows how to estimate size. Multivariate accounting models often require size as an input variable. Also ratios intended to reflect departures from the expected for size could become comparable if their denominator was the same. Such a common deflator would produce easily interpretable residuals.

As seen in section 3.1, log items can be viewed as a unique common effect, size, with some particular variability superimposed. This is true for all the observed items. However, on practical grounds, not all of them are equally adequate to extract size:

- Items like Sales or Current Assets are almost synonymous in log space. Their particular variability is small. They mainly reflect size.
- Inventory, EBIT or Funds Flow have more variability of their own. And items having both positive and negative-valued cases, exhibit a different behaviour in each of such situations. Positive-valued cases are identical to other variates. The negative-valued ones have a very particular behaviour, as far as we could see.
- Finally, Fixed Assets, Working Capital and especially Long Term Debt have large variability of their own. And the non-leveraged firms form a cluster of identical cases.

A proxy for size should therefore be selected from the items mentioned in the first place. However, this proxy would always have, along with the common variability we are interested in, a particular scatter superimposed — the particular variability of the selected item.

How to isolate the common effect? Is it possible to build a variable reflecting only size and having no particular variability of its own? As seen in section 4.1 the common effect isn't directly accessible. However, there is a way of isolating it. We can model the function inverse of ratios. Ratios conceal the common variability in raw data and reveal the particular one. This model conceals their particular variability, thus revealing the common one. Items like Current Assets, Net Worth, Wages and other expenses, and Sales, can be pulled together so as to form one unique variable: If we build, inside each report, geometric means (in log space, averages) of a few items, their particular variability will smooth out and only the common one will remain.

Considering a group of mean-adjusted log items, $e^1, \cdots, e^M$, from the same report, j, we explain their variability as an effect, $\sigma = \log s$, common to them all, plus the deviations from it, ε^i, particular to each item:

$$\begin{array}{lcccl} e_j^1 & = & \log(x_j^1 - \delta_1) - \mu_1 & = & \sigma_j + \varepsilon_j^1 \\ e_j^2 & = & \log(x_j^2 - \delta_2) - \mu_2 & = & \sigma_j + \varepsilon_j^2 \\ \vdots & & & & \vdots \\ e_j^M & = & \log(x_j^M - \delta_M) - \mu_M & = & \sigma_j + \varepsilon_j^M \end{array}$$

The δ_i and μ_i play no active role in this reasoning. They are the thresholds eventually present in the x^i and the expected values of $\log x^i$. We now calculate the average of the $1, \cdots, i, \cdots, M$ mean-adjusted items belonging to the same report:

$$\text{For report } j, \qquad \sigma_j = \frac{1}{M}\sum_{i=1}^{M} e_j^i - \frac{1}{M}(\varepsilon_j^1 + \varepsilon_j^2 + \cdots + \varepsilon_j^M)$$

Since an average of independent random deviates tends to zero with $1/M$, we have for large enough M:

$$\sigma_j \approx \frac{1}{M}\sum_{i=1}^{M} e_j^i \quad \text{ or the equivalent in ordinary space, } \quad s_j \approx \prod_{i=1}^{M}\left[\frac{x_j^i - \delta_i}{\exp \mu_i}\right]^{\frac{1}{M}}$$

Once obtained, s can be used in the denominator of ratios meant to detect deviations from the expected for size. $\sigma = \log s$ is welcome as an input variable in statistical models.

The difficult point here is that the ε^i are not necessarily independent. Some precautions are required before building this model, especially when the variety of available items is limited.

- The items to be used in the building of s shouldn't be correlated. A few items, after deflated by an estimated size, exhibit significant correlation. The introduction of correlated pairs would reinforce any residual variability common to both, instead of smoothing it out.

- The final s shouldn't generate constraints in other items. This is the most difficult condition to achieve. For one reason or another accounting identities seem to propagate across other relations and make themselves present in unexpected situations.

We used two criteria to find the set of items appropriate to build s. The first one is intended to select items. The second one is a test of the applicability of the resulting s.

- After the introduction in the case-average leading to $\log s$ of each new candidate, we compute the variance of this average. If it decreases, the new item is accepted. If it increases, we remove one by one the items already included. For each removed item, if the variance decreases beyond the original value, we replace it by the new one. If the variance never decreases we reject the new item.

- After finding a model for s with minimal variance, we build bivariate scatterplots in which $\log s$ is compared with each one of all the remaining log items in order to find out if constraints or other asymmetry emerge.

	S	NW	W	D	CA	CL	N
1983 VAR:	0.5290	0.5475	0.5473	0.5161	0.5009	0.4999	0.5515
SM:	0.5290	0.5013	0.4971	0.4877	0.4843	0.4800	0.4762
1984 VAR:	0.5807	0.5963	0.6023	0.5349	0.5229	0.5390	0.5972
SM:	0.5807	0.5429	0.5412	0.5251	0.5195	0.5171	0.5138
1985 VAR:	0.5263	0.5591	0.5541	0.4977	0.4829	0.4999	0.5626
SM:	0.5263	0.5032	0.4993	0.4868	0.4812	0.4796	0.4779
1986 VAR:	0.5211	0.5356	0.5419	0.4943	0.5030	0.5009	0.5582
SM:	0.5211	0.4934	0.4920	0.4806	0.4782	0.4771	0.4772
1987 VAR:	0.5318	0.5113	0.5401	0.5003	0.4734	0.4889	0.5646
SM:	0.5318	0.4873	0.4869	0.4795	0.4733	0.4700	0.4716

Table 17: The evolution of the variance of σ for incoming items. VAR is the variance of each log item. SM is the variance of σ, when the items to the left are already in.

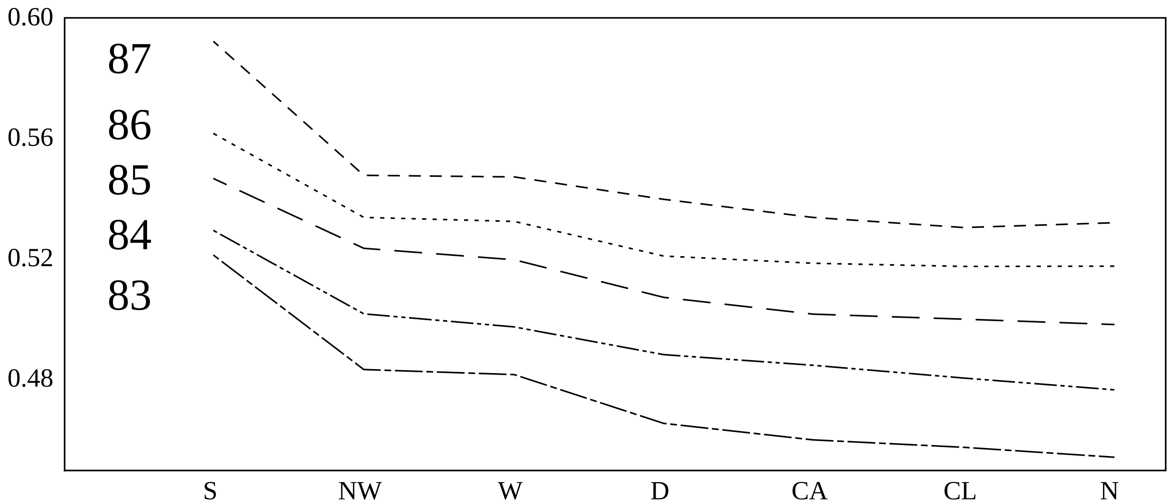

Figure 20: The decrease in variance of σ for incoming items.

For example, in the case of all groups together, the variance of σ decreased whenever the items S, NW, W, D, CA, CL were introduced in the model, for the five samples examined (1983-1987). Other items, (C, I, TC, FA) had the opposite effect. And a few, (EX, N) either made it increase or decrease, depending on the years. Table 17 gathers these results in detail. When reading any row labeled SM from left to right we get a description of the evolution of the variance of σ for an increasing number of items allowed in the case-average. By accepting CA, the variance of this average decreased from 0.4877 to 0.4843 in the 1983 sample. The individual variance of each item is also displayed (rows VAR). Figure 20 on page 63 represents table 17 graphically. It depicts the effect of averaging together more and more items. To avoid the overlapping of the curves the displayed variance suffered a different translation for each year.

Two items emerged as non-adequate to build σ: Inventory and Fixed Assets. Their non-adequacy stems from apportioning more variability than the smooth they produce. Creditors was expected to be non-adequate since it is correlated with Debtors. The same for EX, which is correlated with Sales, and Wages, which is correlated with the Number of Employees. It is indifferent to select one or the other from these pairs, provide both aren't present in the average.

Despite the significant decay in variability obtained, about 14%, none of the above combinations produced exactly symmetrical residuals when deflating items from the Profit and Loss Account. We noticed that, when Total Assets deflates the same items, the asymmetry seems to be smaller than when using s. But, of course, TA performs badly with all the items from the Balance Sheet whilst s will not introduce any asymmetry.

For the set of items we could use here, the best $\sigma = \log s$ seems to be

$$\sigma = \frac{1}{7}\left[\log S + \log NW + \log W + \log N + \log D + \log CA + \log CL\right] \tag{11}$$

In the following, any use of s or σ in this study refers to this particular case-average.

When building models like this one, care must be taken to avoid the accumulation of thresholds. Another problem with this proxy for size is that, if we deflate with s an item already used to build it, the result is the same as if we were using, instead of the entire numerator, a fractional exponent of it. When deflating x_k,

$$\frac{x_k}{s} = \frac{x_k}{\prod_{i=1}^{M} x_i^{\frac{1}{M}}} = \frac{[x_k]^{\frac{M-1}{M}}}{\prod_{i=1}^{k-1} x_i^{\frac{1}{k-1}} \times \prod_{i=k+1}^{M} x_i^{\frac{1}{M-k+1}}} \tag{12}$$

If M is large, $(M-1)/M \approx 1$. But if the number of components of s is small, the exponent affecting the numerator models a non-linear relation. Therefore, it would be interesting to find a large number of items to build s, also because the self-smoothing would improve. Anyway, the set displayed in equation 11 performs remarkably well.

The distribution of size: $\sigma = \log s$ has a much smaller kurtosis than that observed in the case of the logs of raw numbers. The reason for this is straightforward. The variability of σ is the one of the main axis of a multivariate distribution, the one which is the source of the Gaussian behaviour of log data. This reasoning was introduced in chapter 4. We recall figure 19 (page 58). In this graphical representation, the variability of σ would be the one along the "Size Axis". The source of positive kurtosis is the "Ratio Axis". Log items are 45^o projections of this multivariate distribution. That's why they contain some kurtosis.

Notice that ratios with s in the denominator no longer yield contrasts between two departures from size. Ideally, they reflect the real departure from size of the item in the numerator. Using our notation, we can now access each f^x or ε^x. As a consequence, we can also use size-adjusted Sales, Working Capital or Debt, separately as input variables in statistical models. Such variables are self-explanatory to an extent so far unattained. Their interpretation is immediate, and it is expected that the indiscriminate use of all sorts of ratios in multivariate statistical models, could be avoided.

6.2 The Homogeneity of Industrial Groups

In the presence of groups, the modelling of accounting relations cannot avoid two important questions:

- Is a particular grouping significant so that it should be taken into account? If the data is more similar inside groups than from group to group this is the case.
- Are groups similar in their effects upon the features of the data? For example, is liquidity affected in the same way as, say, profitability in the presence of an industrial grouping?

In this section we show how to answer the first question, using the industrial grouping as an example. We compare the variability inside industries with the one between them. As a result we obtain a measure of the importance of this grouping, for each variable involved. The second question can be answered at several levels of accuracy. Here, we describe the simplest procedure, consisting of ranking a measure of homogeneity by industry, and then verifying if these rankings are consistent for different variables.

6.2.1 Introduction and Related Research

Accounting reports don't contain all the information necessary to characterize firms. The very basic problem of financial statement analysis is the existence of similar accounting patterns which are not neighbours in the space of the real features of firms. In order to correctly map firm features, accounting data isn't enough. External information is also required.

A clear example of this, is the industrial classification. The similarity of firms as perceived by the SEIC can be different from the similarity of accounting reports. A non-standard piece of information, the number of employees, turns out to be important when checking the homogeneity of industrial groups. Other non-accounting variables, eventually also important, could be the patterns of consumption of energy, area requirements for plant or stores, the age of the firm, and its location.

The use of a limited amount of information like the one contained in accounting reports, generates extra unexplained variability in statistical models. Here, we are not concerned with the amount of this extra variability. We are interested in its complexity: The complexity in models becomes higher when the data reflects facts we cannot account for. For example, the Leather or the Wool industries could add complexity to the model if the accounting features of each firm were influenced by its location. Location would act as a hidden grouping.

Fixed and random effects: Some groupings are defined a-priori by an accepted institution like the SEIC in the UK. Others are the result of objective causes. The grouping of firms into failed and non-failed has a statistical nature which is different from the SEIC grouping. The former introduces in the population a simple partition. The later introduces real variability. Simple partitions are known as fixed effects. Groupings which introduce randomness are known as random effects.

Groupings that introduce random effects in a population can indeed introduce more than one. The assessment of the number of independent sources of variability a grouping carries with it, is eventually important. For example, if a particular grouping contains two random effects, it is likely to induce higher order relations between input variables, thus requiring non-linear tools to be modelled.

Related research: Firm grouping is itself not a very homogeneous body of research. It includes simple industry comparisons of ratios, tests on widely accepted groupings of firms and the search for clusters of firms according to similarities of ratios and other data. The former topic has been explored from very early in the literature. Foster [15] offers an overview. There is an established evidence on differences between some ratios for well known industry groups.

The search for clusters of firms has been carried out by Elton and Gruber [12], and Jensen [21] amongst others. Sudarsanam and Taffler [29] (1985) tested the separability of the SEIC groups, using accounting information.

6.2.2 Measuring the Significance of a Grouping

We are interested in assessing the extent to which the SEIC grouping of industries is effective in creating more similar subsets of firms. Given that the 14 industrial groups selected in this study represent a

Item	1983		1984		1985		1986		1987	
	F	ρ	F	ρ	F	ρ	F	ρ	F	ρ
SIZE	5.34	9.87%	4.42	6.88%	4.48	6.70%	4.86	7.15%	4.87	7.31%
S	7.21	13.53%	6.61	10.79%	6.53	10.27%	6.48	9.85%	5.93	9.11%
NW	4.00	7.09%	3.94	6.00%	3.63	5.23%	4.20	6.10%	4.17	6.15%
W	5.01	9.26%	3.51	5.17%	4.09	6.03%	4.52	6.59%	4.66	6.96%
I	3.31	5.56%	2.85	3.87%	3.03	4.08%	3.12	4.12%	2.99	3.98%
D	7.02	13.20%	6.51	10.65%	6.31	9.91%	6.47	9.85%	6.39	9.92%
C	5.68	10.58%	5.17	8.26%	5.22	8.03%	5.48	8.22%	5.45	8.35%
FA	6.45	12.08%	6.29	10.23%	6.52	10.26%	7.18	11.00%	6.65	10.35%
CA	4.62	8.42%	4.10	6.32%	3.80	5.51%	3.86	5.44%	3.60	5.07%
CL	5.42	10.03%	5.14	8.20%	4.97	7.59%	5.67	8.52%	5.06	7.65%
N	4.41	8.00%	3.55	5.24%	3.93	5.76%	4.06	5.77%	4.08	5.94%
EBIT	5.88	11.73%	4.46	7.41%	3.49	5.26%	3.74	5.61%	5.45	8.86%
FL	5.90	11.51%	4.75	7.76%	4.35	6.76%	4.36	6.60%	4.94	7.82%
DEBT	4.44	11.86%	3.04	6.10%	3.84	7.66%	3.54	6.43%	4.08	7.81%

Table 18: The F statistic and the Intra-Class correlation, ρ, when log items were used to explain the industrial grouping.

sampling amongst a larger number of choices, it is inappropriate to use fixed-effects models. Thus, we use a random-effects one, the intra-class correlation.

The intra-class correlation coefficient, ρ, measures the proportion of the total variability that is associated with a grouping. It is a standardized way of comparing the variability within groups with the one between groups, when the effects are random. If s_b^2 is the expected value of the mean-squares between groups and s^2 is the corresponding mean-squares within groups, an estimator of ρ is

$$r = \frac{s_b^2 - s^2}{s_b^2 + (k-1) \times s^2}$$

k is the number of cases in each group. For M groups of unequal size $n_i, i = 1, M$ and $N = \sum n_i$, k should be approximated as

$$k = \frac{1}{M-1} \times \left(N - \frac{\sum n_i^2}{N} \right)$$

It is possible to estimate confidence intervals for r. A detailed discussion of this statistic and the way it is derived can to be found in Snedecor and Cochran [27] (pp. 242 in the 9^{th} ed.).

The more similar the groups are, the more the correlation intra-classes approaches 1. When the variability inside groups is smaller than the one in the whole sample, this measure yields a positive value. For a variability inside groups similar to the one between them, the intra-class correlation yields zero. In the case of groups containing more variability than the whole, a negative correlation emerges. When the effects governing the variability within groups and between them are independent, negative ρ cannot occur. Negative ρ emerge only in cases where the effects interact.

The data: During the usual period of five years we examined three kinds of accounting information. Firstly, several log items and also s, our estimated size. Secondly, the logs of the same items after deflated by s. Finally, the logs of a few ratios.

In tables 18, 19 and 20 on pages 66 and next, we display the estimated intra-class correlation along with the F statistic. The number of firms involved ranges from 555 to 702 in 14 industries.

Results: Raw numbers. The logs of raw data show a small but significant increase in homogeneity owing to the industrial grouping. The values of ρ are stable during the considered period and no

Residual	1983		1984		1985		1986		1987	
	F	ρ	F	ρ	F	ρ	F	ρ	F	ρ
S	11.79	21.41%	15.27	23.54%	16.83	24.64%	13.94	20.53%	11.39	17.45%
NW	1.54	1.35%	1.92	1.96%	1.69	1.43%	1.71	1.42%	1.61	1.25%
W	5.28	9.84%	6.18	10.12%	6.34	10.00%	6.85	10.50%	6.97	10.87%
I	4.36	7.88%	4.87	7.77%	4.35	6.57%	5.36	8.10%	3.35	4.65%
D	8.57	16.06%	11.05	17.85%	9.40	14.82%	7.18	10.98%	6.06	9.36%
C	3.46	5.85%	3.45	5.03%	2.86	3.71%	2.77	3.42%	3.15	4.22%
FA	2.23	3.00%	2.87	3.89%	3.35	4.65%	4.00	5.67%	3.69	5.20%
CA	2.79	4.36%	3.87	5.85%	4.11	6.08%	4.02	5.72%	2.04	2.09%
CL	2.93	4.65%	3.94	5.96%	2.67	3.34%	2.32	2.57%	1.69	1.40%
N	9.75	18.24%	13.42	21.23%	12.57	19.42%	12.10	18.17%	9.46	14.76%
EBIT	6.73	13.49%	5.16	8.77%	3.30	4.87%	2.98	4.11%	3.35	4.89%
FL	8.13	15.91%	6.13	10.29%	4.55	7.13%	2.79	3.62%	3.65	5.40%
DEBT	1.52	1.99%	0.97	-0.08%	1.85	2.41%	1.59	1.56%	1.91	2.44%

Table 19: The F statistic and the Intra-Class correlation, ρ, when several different log residuals were used to explain the industrial grouping.

negative or zero cases were observed. They aren't different from one another, as expected. In fact, since the logs of raw numbers mainly reflect size, they yield similar proportions of variability associated with grouping. Fixed Assets, Debtors and Sales are the most homogeneous log items inside groups (10%). Inventory is the least homogeneous (4%). Size itself is similar to many other items (5%). On the whole, the homogeneity ranges between the extreme values of 3% and 13%.

Results: Size-Adjusted. The contrast between industries increases for size-adjusted items (the ε^x). Some of these deviations from the expected for size show a much larger homogeneity intra-groups than others. Also, the consistency for the considered period of five years is not affected in most of the items but it is completely lost in a few. Gross Funds from Operations and EBIT, for example, plunge from a strong similarity inside industries to a much smaller one from 1986 on. It seems as if profitability were increasingly non-homogeneous. See table 19 on page 67.

Size-adjusted Sales and the Number of Employees are the most similar inside groups. The SEIC seems to rely on these items as a criterion for determining groups. Next, Debtors and Wages. Debt and Net Worth are the less homogeneous: The financial structure of firms isn't sensitive to industrial groups. On the whole, the homogeneity of the residuals ranges from naught to 25%. These values denote a more diversified influence of the industrial grouping upon size-adjusted features than upon raw data.

Results: A few ratios. Table 20 on page 68 displays the intra-class correlations for a few more ratios. The above size-adjusted measures are also ratios, of course: They capture deviations from the expected for size. The Long Term Debt to Net Worth ratio shows no traces of recognizing the SEIC grouping as such. The liquidity ratio yields measures of similarity comparable with those of raw data. Ratios incorporating Sales, Wages, Debtors or the Number of Employees, clearly recognize the SEIC grouping. If our goal were the identification of ratios appropriate to model the SEIC, then the W/N ratio would be a good choice.

A method to select appropriate ratios for specific tasks could consist of using ρ. Firstly, the intra-class correlations of many size-adjusted log items would be assessed. Then, the most promising combinations of items would be selected amongst the residuals with highest ρ, and tested.

Residual	1983		1984		1985		1986		1987	
	F	ρ	F	ρ	F	ρ	F	ρ	F	ρ
S/N	10.08	18.80%	14.22	22.30%	13.44	20.57%	12.84	19.14%	11.08	17.11%
$(S \times N)/s^2$	11.50	21.14%	15.75	24.24%	16.47	24.39%	13.70	20.25%	9.98	15.52%
W/N	11.14	20.62%	16.67	25.42%	18.43	26.62%	20.25	27.86%	16.09	23.61%
CA/CL	3.71	6.44%	4.76	7.56%	3.86	5.63%	5.04	7.51%	2.56	3.11%
$DEBT/NW$	1.65	2.51%	0.97	-0.10%	2.07	3.08%	1.73	1.96%	1.90	2.45%
$S/EBIT$	8.00	16.02%	6.98	12.14%	8.69	14.62%	6.67	10.97%	5.70	9.32%

Table 20: The F statistic and the Intra-Class correlation, ρ, when a few log ratios were used to explain the industrial grouping.

Summary: The industrial grouping clearly gathers firms which are, to a small extent, more similar regarding size. Also, a few features of the firm are more homogeneous inside industries. It is the case for Sales, Wages, the number of employees or Debtors. The financial structure of firms is not especially more similar inside industries, and the measures of profitability seem to yield very different results from year to year. In the early years of our observations the profitability of firms is remarkably similar inside the same industry. In the later ones (1986, 1987) it becomes irregular.

There is nothing in the obtained results able to defy the common-sense of accounting knowledge. The results are expected. A very simple technique yielded consistent and interpretable results.

6.2.3 Assessing the Complexity of Grouping Effects

This section tries to answer the second question posed above. We are interested in broadly knowing if it is acceptable to consider one unique random effect in the SEIC grouping. It isn't a particularly interesting subject for fields other than the multivariate modelling of accounting relations.

The method: Building maps from distances. As remarked before, this problem can be treated with different levels of accuracy. Here we selected a simple one. It could be much improved, so that we would end up with a real complex instrument to measure complexity.

Our method is based on the well known possibility of constructing maps from distances. For example, it is possible to build a map showing the relative positions of the main cities in Britain just by knowing the distances between them. Cities require two dimensions to be mapped. When the objects to be mapped lie in a straight line the result of this building can be expressed, if desired, as a simple ranking. Objects positioned so as to form a two-dimensional map cannot be ranked.

We are interested in discovering if it is acceptable to rank the industrial groups according to the variability of accounting features. If it turns out that the different industries can be ranked according to this variability, then the SEIC grouping is likely to introduce just one random effect. On the contrary, if the variability introduced by the SEIC resists a simple ranking — thus requiring a two-dimensional map like in the case of cities — then the variability introduced by the SEIC is complex.

In fact, if the SEIC grouping is a unique effect, it will impinge upon the features of the firm in different degrees but not in different directions, thus yielding a consistent variability for several features. For example, if Chemicals has smaller variability in liquidity than Food, then Chemicals would also exhibit a smaller variability in profitability or any other feature. But if, in the former group, there is a smaller variability in liquidity when compared with Food, and a larger one for profitability, higher order (complex) effects are expected.

The method we developed to test the complexity of grouping consists of:

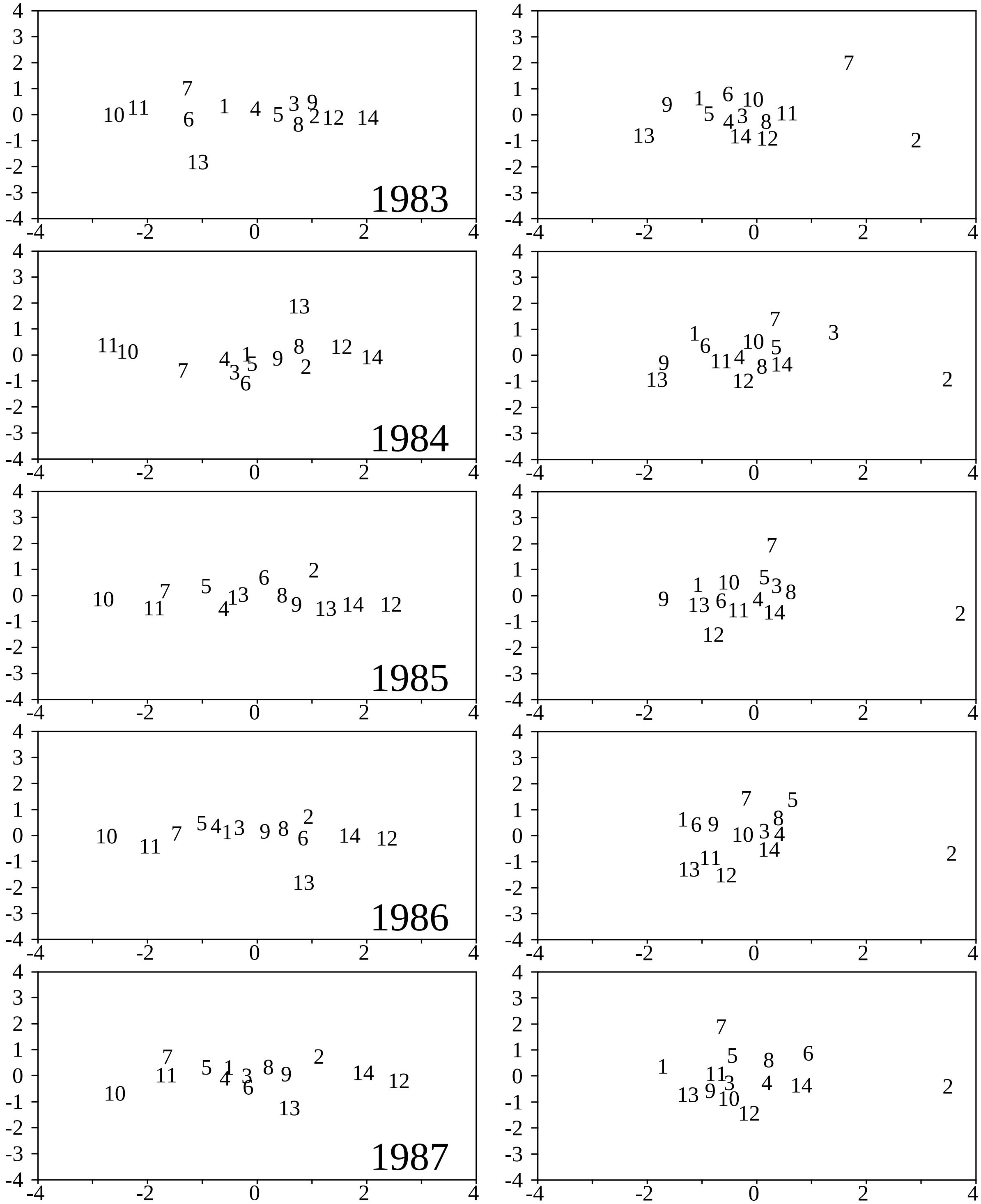

Figure 21: Each one of these scatterplots is a two-dimensional map showing the position of industries regarding the variability of their accounting information. The X-axis is the first dimension. The Y-axis is the second one. On the left, logs of raw data during five years. On the right, log deviations from that expected for size during the same period.

1983	1984	1985	1986	1987
CLOTHING	WOOL	CLOTHING	CLOTHING	CLOTHING
WOOL	CLOTHING	WOOL	WOOL	MACHINE TOOLS
LEATHER	MACHINE TOOLS	MACHINE TOOLS	MACHINE TOOLS	WOOL
MACHINE TOOLS	CHEMICALS	ELECTRICITY	ELECTRICITY	ELECTRICITY
INDUSTRIAL PL.	INDUSTRIAL PL.	CHEMICALS	CHEMICALS	CHEMICALS
BUILDING MT.	PAPER & PACK	PAPER & PACK	BUILDING MT.	BUILDING MT.
CHEMICALS	ELECTRICITY	BUILDING MT.	PAPER & PACK	INDUSTRIAL PL.
ELECTRICITY	BUILDING MT.	INDUSTRIAL PL.	MOTOR COMPON.	PAPER & PACK
PAPER & PACK	ELECTRONICS	ELECTRONICS	ELECTRONICS	ELECTRONICS
ELECTRONICS	MOTOR COMPON.	MOTOR COMPON.	INDUSTRIAL PL.	MOTOR COMPON.
MOTOR COMPON.	METALLURGY	METALLURGY	METALLURGY	LEATHER
METALLURGY	TEXTILES M.	LEATHER	LEATHER	METALLURGY
TEXTILES M.	FOOD MANUF.	FOOD MANUF.	FOOD MANUF.	FOOD MANUF.
FOOD MANUF.	LEATHER	TEXTILES M.	TEXTILES M.	TEXTILES M.

Table 21: Industries ranked by variability of log items. Below, the largest variability.

- Firstly, the standard deviation of several features of the firm is measured for a sampling of groups.
- Secondly, joint distances between industrial groups are computed from these standard deviations. One typical such distance could be the Euclidean one.
- Finally, ordinal scores are discovered that position each industry according to the above distances.

The final result is a map. Each industry is a position in that map. The coordinates are the obtained scores. If industries lie in a straight line, dimensions other than the first one are negligible. This means that groups can be ranked according to a unique measure of variability. They are affected by the grouping in different degrees but not in different directions. On the contrary, if two dimensions are required to position groups, they cannot be correctly ranked using a joint measure of variability. Our method also points out which industries are likely to be contributing to an increase in complexity.

The data: We used two sets of data. Firstly, the standard deviations of all the log items used in previous section. Secondly, the standard deviations of log deviations from that expected for size (the ε^x). The first set is a comparison term. As the logs of raw data mainly reflect size, their complexity should be small. The results obtained for the second set can be compared with these. The experiment was carried out for the usual period of five years.

Results: Log items. Figure 21 shows the obtained two-dimensional maps. The maps on the left refer to raw data. The ones on the right, to size-adjusted items. The X-axis of each map is the first dimension and the Y-axis is the second one. Inside maps, the position of one industry is highlight by a number. The correspondence between these numbers and industries can be found in appendix, in table 34 on page 100.

Maps on the left are clearly different from those on the right side. In the case of raw data there is a trend towards a straight line, as the effect of size is preeminent. Industries have different variability but they are under the same effect. Leather is the exception. It shows signs of influences other than size, during three of the observed years. The most homogeneous industries regarding size are Clothing, Wool, and Machine Tools. The least homogeneous are Miscellaneous Textiles, Metallurgy and Food Manufacturers. We show these ranks in table 21 (page 70).

Industries like Building Materials, Metallurgy, Machine Tools, Clothing and Food show a consistent variability during the whole period. Chemicals, Electricity, Electronics, Motor Components and Wool are also regular. Paper and Packing, Industrial Plants, and especially Leather, are irregular. Their ranking is not consistent during the whole period.

Results: Size-adjusted items. The two-dimensional maps on the right side of figure 21 show clear differences from those on the left. The trend towards one unique dimension is no longer visible. One industry, Metallurgy, emerges as very particular, with a larger variability than the others. We recall from section 3.1 that Metallurgy was also unique in the proportion of variability of log items that size would explain. The scores obtained are consistent for the five years.

Leather, Motor Components and Building Materials are now the most homogeneous groups. Metallurgy is the least. Building Materials, Chemicals, Electronics, Clothing, and Food, exhibit the same score during the whole period. This means a persistent variability associated with internal features. Industries like Electricity, Metallurgy, Motor Components, Wool, and Leather, also show a reasonable stability. Paper and Packing, Industrial Plants, and especially Machine Tools, aren't stable in their homogeneity. Three industries emerge as showing a consistent behaviour. Both the variability of size measures and features, determine a clear position in the case of Food, Clothing, and Building Materials. Industrial Plants, Paper and Packing, and Leather, are examples of the opposite behaviour. The remaining industries lie in between.

6.3 Summary

In this chapter we explored the main sources of variability of accounting data. Firstly we produced a set of statements for guidance in the search for a proxy for size. We have shown that simple case-averages of the logs of selected items produce a significant reduction in the total variability and can be used to isolate the statistical effect of size.

If we use this estimated size as the denominator of ratios, we obtain a variable that enhances the interpretability of results in statistical models and avoids the excessive number of input variables.

We also studied the importance and effect of the SEIC industrial grouping when present in accounting relations. Our results show that both the variability of size and the one of some features of firms are dependent on grouping. But in the last case, the effect of grouping is not similar across industries.

The existence of higher order effects questions the use of linear tools to model accounting relations.

Chapter 7

The Complement of a Ratio

Ratios only use part of the information needed to build them. Given two items, y and x, is there any interesting piece of information besides the one conveyed by the ratio y/x? In this chapter we show that the size-adjusted information contained in two items can be expressed in terms of the ratio itself plus a remainder. We also point out, based on the way some ratios are used in practice, that such a remainder is likely to be valuable for financial analysis.

We introduced earlier in this study the notation $\varepsilon_j^{y/x}$, or the corresponding $f_j^{y/x}$ in ordinary space, to designate the deviation of the ratio y/x from the standard, in the case of firm j. We also used the ε_j^x, or the f_j^x, to designate the deviation of x from the expected for the size of firm j. Such a deviation is the contribution of firm j to the variability particular to x.

7.1 Assessing the Remainder

Since any pair of items, $\{x, y\}$, conveys two-dimensional information, and ratios are just one variable, when we use ratios instead of their components we put aside information. Not only the one about size. We put aside size-adjusted information, potentially interesting on grounds of financial analysis. The size-adjusted information conveyed by the ratio y/x can be written in log space as the subtraction of two deviations from the expected for the size of the firm. If $y/x = R_{y/x} \times f^{y/x}$, $R_{y/x}$ being the standard for that ratio, then

$$\log f^{y/x} = \varepsilon^y - \varepsilon^x, \qquad \text{in which} \quad \left\{ \begin{array}{lcl} \varepsilon^y & = & \log(y - \delta_y) - \mu_y - \sigma \\ \varepsilon^x & = & \log(x - \delta_x) - \mu_x - \sigma \end{array} \right.$$

Given that it is possible to find an estimated $\sigma = \log s$, (see section 6.1) we can now isolate both the size-adjusted information contained in x, the ε^x, and the one of y, the ε^y.

Let us define two Cartesian coordinates in which the ε^y are measured along the Y-axis, and the ε^x along the X-axis. All the size-adjusted information conveyed by y and x about firm j will be represented by a point, $\{\varepsilon^y, \varepsilon^x\}_j$, in this coordinate system. Now we rotate this system 45^o anti-clockwise. We obtain new coordinates in which the X-axis measures $\varepsilon^y - \varepsilon^x$ and the Y-axis measures $\varepsilon^y + \varepsilon^x$. But $\varepsilon^y - \varepsilon^x$ is the information conveyed by the ratio y/x. Since the new Y-axis is orthogonal to the one assessing the information conveyed by the ratio, we can be sure that all the information not accounted for by the ratio itself will be contained in $\varepsilon^y + \varepsilon^x$. Hence, $\varepsilon^y + \varepsilon^x$ conveys the size-adjusted information contained in x and y, but not contemplated by y/x.

It is easy to see that, in ordinary space, $\varepsilon^y + \varepsilon^x$ is the ratio $(x \times y)/s^2$, s being the estimated size. Therefore, given any ratio y/x, if we want to know which size-adjusted information conveyed by its components hasn't been assessed by it, we should look into the ratio $(x \times y)/s^2$. The ratio $(x \times y)/s^2$ is a complementary ratio to y/x. Together, they describe the two orthogonal aspects of a unique two-dimensional observation.

7.2 Is the Remaining Information Useful?

In this section we show that the pair of complementary variables obtained above can answer two kinds of questions that specific pairs of ratios are meant to answer as well. From this, we conclude that those pairs convey, in some cases, information useful to analysts. We also discuss the cases in which it is likely to find complementary information that is useful.

It is often mentioned in the literature that ratios are used because of the need to control for size. This control for size has several meanings. Two of them are now considered in the form of questions:

- Is a particular item big (or small) when compared with the size of the firm? This is the problem of assessing deviations from standards for size. Financial ratios are meant to answer this question when the deflator is selected so as to reflect size (Total Assets is a typical choice). In our framework the answer to the above question is given by the ratio $(x \times y)/s^2$ (in log space, the $\varepsilon^y + \varepsilon^x$).

- To what extent a given feature of the firm, like liquidity, is far away from the expected, regardless of the magnitude of its components when compared with the standards for size? This is the problem of measuring departures from standards describing features by themselves. In such cases the deflator is selected so as to produce a contrast when compared with the deflated item. In our framework this would be accomplished by the ratio itself (the y/x or, in log space, the $\varepsilon^y - \varepsilon^x$).

Though no sharp separation exists between both functions, some ratios seem more intended to answer the first question whilst others are more intended to answer the second one. For example, in the two ratios Working Capital to Total Assets and Current Assets to Current Liabilities, the first one assesses liquidity by referring it to the size of the firm, whilst the second one assesses the feature emerging when contrasting short term assets with liabilities, regardless of the size of the firm.

In many other situations, ratios are used alone or their pairing isn't related to the problem discussed here. For example, the Interest Cover ratio, despite being often used along with Financial Structure ratios, doesn't relate to them in the way we discuss here. Analysts seek two pieces of information which are complementary on grounds of financial analysis, not because of any complementary relation based on information content.

Applicability of the complementary ratios: Discussion and related research. As stated above, in many cases the remaining information isn't used by analysts. We now discuss the cases in which pairs of complementary ratios could yield interesting information. The literature concerning this topic is scarce. A subject somehow related is Horrigan's response to Barnes [19]. Horrigan claims that the main task ratios undertake is the assessment of specified relationships. "They adjust for the data size effect only incidentally. (...) Size deflation is certainly an interesting property of financial ratios, but it is hardly their major purpose." Horrigan seems to suggest that only the $\varepsilon^y - \varepsilon^x$, the y/x, convey interesting information. However, on statistical grounds, it is easy to sustain the usefulness of the

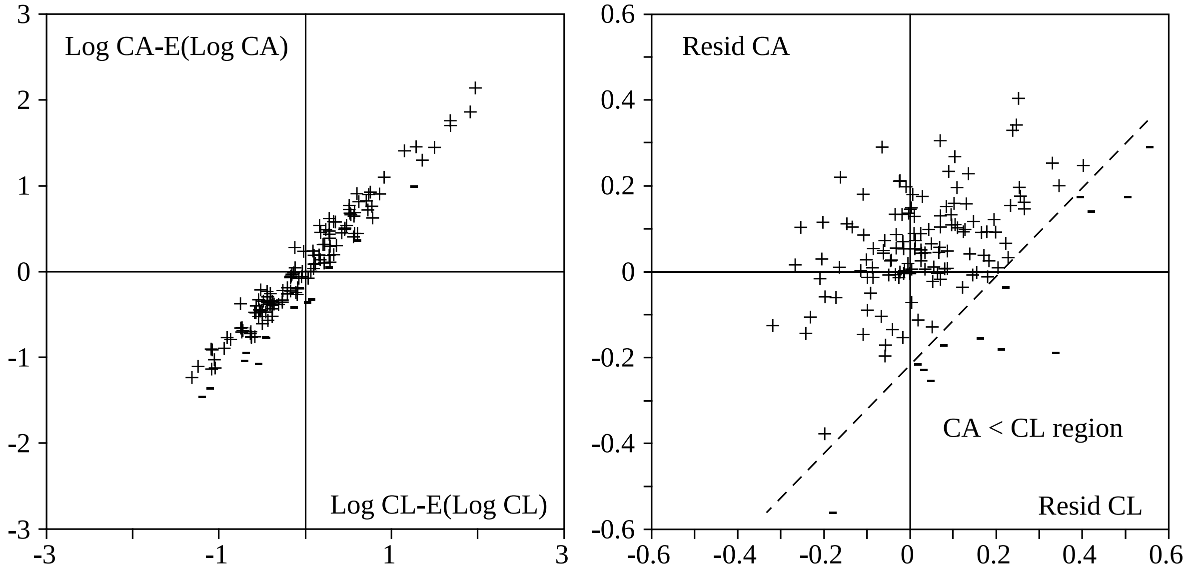

Figure 22: On the left, a mean-adjusted scatterplot of CA (Y-axis) versus CL (X-axis). On the right, the corresponding residual plot. Electronics, 1986. The negative sign means $CA < CL$.

complementary ratio (the $(x \times y)/s^2$), since it is orthogonal to y/x. Only when ε^y is correlated with ε^x the above decomposition of information becomes less attractive. This is because correlation means redundancy: On the limit, two strongly correlated ε^y and ε^x would carry the same information.

Uniqueness of bivariate information: We now explore a more specific question. When the remaining information turns out to be useful, to what extent is it convenient to gather in a unique observation the two ratios — the y/x and the $(x \times y)/s^2$ — instead of using each of them separately? Is there anything to be gained by using bivariate information instead of their two separated pieces?

The bivariate information conveyed by the pair of complementary ratios $\{y/x, (x \times y)/s^2\}$ can be more revealing than the examination of the two ratios separately. Firstly, because two dimensions allow an increasing in specificity: Trajectories, recognized as such, are more accurate and easy to interpret than trends. Secondly, because, in some cases, the bivariate information could be unique. This happens whenever the scatter of cases draws, in two-dimensions, a shape impossible to describe functionally. For example, when the scatter of cases is less dense in one quadrant than in the other three, or when there is a comet-like shape (a bivariate tail).

That's why, in the examples to be explored in next section, we privilege bivariate representations (scatterplots) instead of studying both components separately.

7.3 Creating Bivariate Tools

In this section we present increasingly elaborated scatterplots of accounting data leading to what we call the Rotated Residual Plot. The Rotated Residual Plot is a bivariate representation of the two complementary ratios studied in previous section.

7.3.1 Non-Rotated Plots

Simple visual inspection of two items can be achieved with scatterplots. For the ratio y/x, $\log x$ would be the abscissa and $\log y$ the ordinate. Figure 5 on page 21 contains one of these "XY Plots" as we call them. Scatterplots in log space reveal the existence of external forces and significant thresholds. For example, in the case of constraints introduced by accounting identities, all cases gather on one side of the line $y = x$, as seen in figure 16 (on page 50). Non-negligible thresholds determine non-linearity in the relation between $\log x$ and $\log y$. Ratios like $EBIT/S$ or CA/FA exhibit, in a few industrial groups, traces of non-linearity consistent with this hypothesis.

The information conveyed by the XY plot contains the one of ratios. The horizontal (or vertical) distances from any case to the line $\log y - \mu_y = \log x - \mu_x$, which is the axis with largest variability, measure deviations from ratio standards. For example, the scatterplot formed with $\log CA$ in the abscissa and $\log CL$ in the ordinate, yields, for any point $\{CA_j, CL_j\}$ representing the position of firm j, a measure of $\varepsilon_j^{CA/CL} = (\log CA_j - \overline{\log CA}) - (\log CL_j - \overline{\log CL})$ which is the deviation from the ratio standard in log space. As usual, $\overline{\log x}$ stands for the log median (the mean of $\log x$).

The mean-adjustment: A first step towards more practical tools is the mean-adjustment of the data. Financial diagnosis is based on the magnitude of deviations from standards. The value of the standard itself is important only in that it allows the calculation of such deviations. Therefore, the mean-adjustment throws away a non-important piece of information. Mean-adjustment is also useful when it is convenient to gather in the same scatterplot data belonging to several years. In this case we mean-adjust separately each year.

The residual plot: If we plot, instead of mean-adjusted items, the residuals obtained after controlling for s, the common effect, we get a scatterplot of ε^y with ε^x. This residual plot is adequate to detect correlations between features of the firm. Since the strong effect of size has been accounted for, any residual correlation becomes visible. XY plots aren't accurate in detecting residual correlations since the effect of size, having a much larger variability, completely masks them. Figure 22 compares a mean-adjusted plot (left) with a residual one for the same data (right).

7.3.2 Rotated Plots

The two plots presented next are intended for financial analysis. The first one, we call the rotated plot, preserves information regarding the size of the firm. The second one, we call the rotated residual plot, only shows size-adjusted information.

Taking size into account: The rotated plot. Given a data matrix $X_{N\times 2}$ containing N cases of two mean-adjusted log items, $\log y_{\text{adj}}$ and $\log x_{\text{adj}}$, we obtain the corresponding 45^o anti-clockwise rotated data matrix $X^r_{N\times 2}$ by applying the transformation H as in

$$X^r = XH \quad \text{with} \quad H = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{13}$$

The resulting variables are $h_1 = \log y_{\text{adj}} + \log x_{\text{adj}}$ and $h_2 = \log y_{\text{adj}} - \log x_{\text{adj}}$. h_1 is size-preserving and h_2 is, in log space, the deviation of y/x from the standard. We can build a scatterplot in which h_1

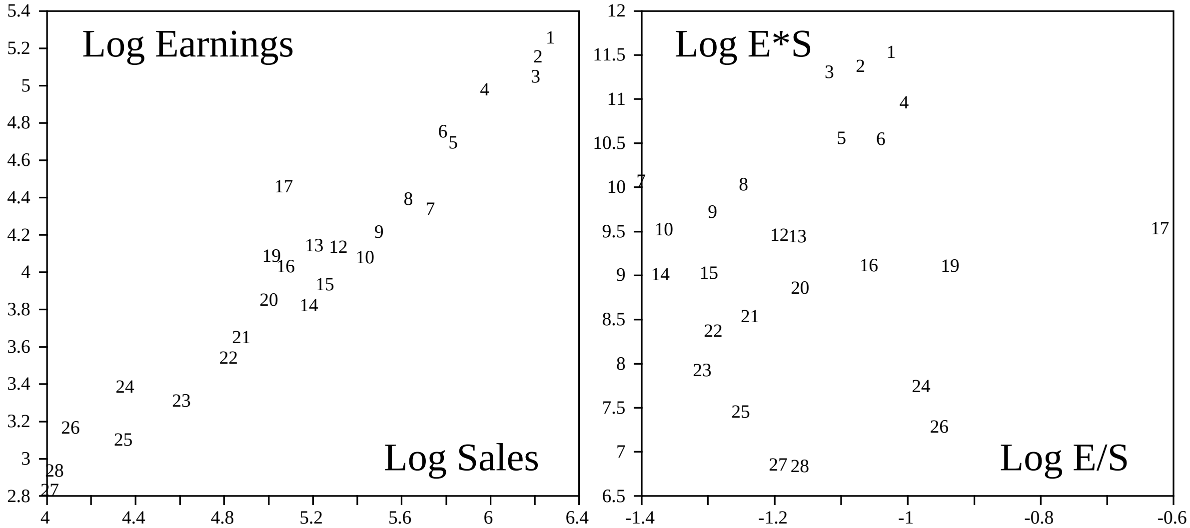

Figure 23: On the left, a XY plot of Earnings (Y-axis) with Sales (X-axis). On the right, the corresponding rotated plot. Motor Components, 1984.

is the ordinate and h_2 is the abscissa. The space spanned by $\{h_1, h_2\}$ contains the same information conveyed by a mean-adjusted XY plot, but now arranged in a way that makes sense in financial analysis.

This rotated plot allows the straightforward measuring of deviations from ratio standards along the X-axis. It is more accurate than the XY plot since the cases now span more uniformly the whole of the neighbourhood about their expected values. Figure 23 shows, on the left, the usual XY plot and on the right the same data after rotation.

Rotated plots like the one of figure 23 (right) retain in the Y-axis the information regarding the size of the firm. They are ideal for observing features described both by a ratio and by size. In the literature there is a growing conscience about the importance of size — not just deviations from the expected for size — in some specific problems. For example, when predicting firm distress, both the ratio Cash-Flow to Total Debt and the size of the firm seem to be revealing. We can put together both pieces of information by means of this plot.

The rotated residual plot: In the rotated plot, the coordinates of any point, $\{h_1, h_2\}$, are ratio residuals and joint measures of size. For the tracing of features unrelated to size the rotated residual plot is the adequate tool. It uses, instead of size, a ratio which is the information-complement of the one the X-axis shows. In this rotated residual plot:

- The X-axis measures, in log space, the deviation of the ratio y/x from the standard. For conveniently selected y and x, this axis is supposed to capture a financial feature of the firm.
- The Y-axis measures the deviation of $(x \times y)/s^2$ from the expected, that is, the joint departure of y and x from the expected for that size.

Notice that the rotated residual plot is just a 45^o anti-clockwise rotation of a residual plot. In a residual plot, ε^y is the abscissa and ε^x is the ordinate. After rotating these axis in the way described in (13) we obtain a new X-axis assessing $\varepsilon^y - \varepsilon^x$, which is the residual of y/x, and a new Y-axis assessing $\varepsilon^y - \varepsilon^x$ which is the residual of $x \times y)/s^2$.

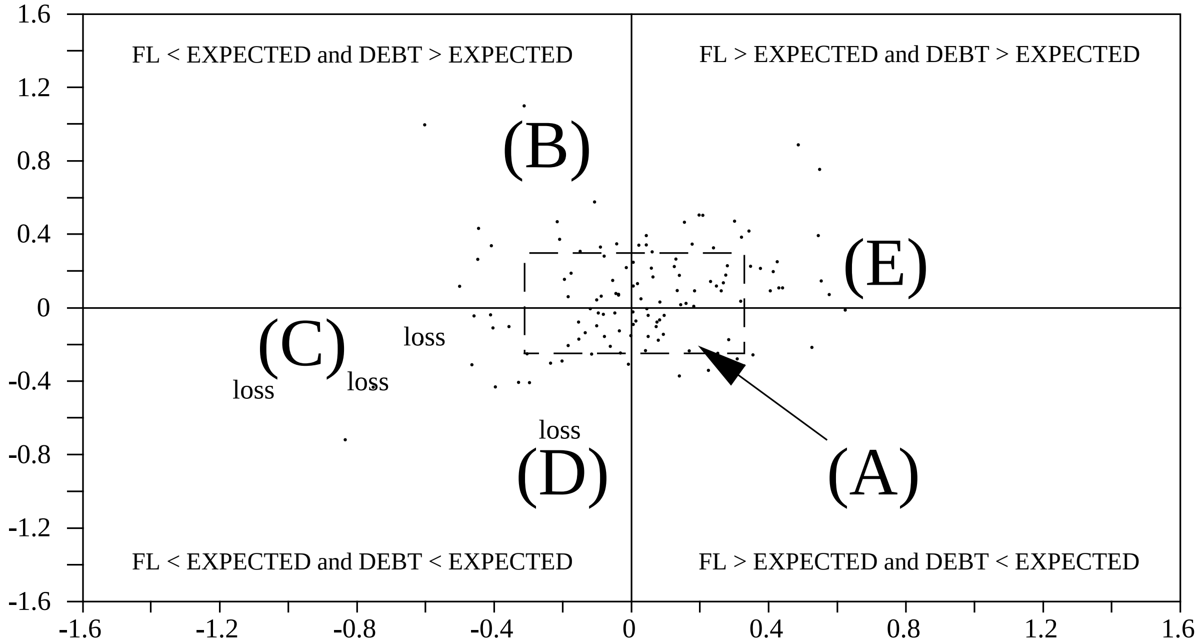

Figure 24: Diagnosis and location of firms in the rotated residual plot. Funds Flow to Total Debt.

The rotated residual plot can also be viewed as a way of reducing from three to two the dimensions of the information concerning a pair of items. In fact, the whole of the potentially interesting information involving two items, y and x, is a three-dimensional vector, $\{\varepsilon_y, \varepsilon_x, s\}$. For example, when measuring liquidity, we might hesitate between using CA/CL, thus getting a picture of liquidity by itself, or using WC/TA in order to have insight into the position of liquidity regarding the size of the firm. The three-dimensionality of the desired information is depicted by the fact that, only when knowing CA, CL and TA, would we be able to answer both questions. A rotated residual plot could assess liquidity, both by itself (along the X-axis), and as referred to the size of the firm (Y-axis). The dimension reduction has been achieved by pulling together, along the Y-axis, the departures from size observed on y and x.

7.3.3 Financial Diagnostics and the Rotated Residual Plot

Figure 24 shows the diagnostics to infer from the location of any firm in the rotated residual plot. As usual, the term "feature" refers to characteristics of firms reflected by financial statements: Liquidity, profitability, financial structure and so on. Ratios are supposed to capture features. In this sense, the diagnostics provided by the rotated residual plot are:

The position of a firm is A: Both the feature we are interested in and its magnitude regarding the size of the firm, are near the expected for that industry.

The position of a firm is B: The feature we are interested in is near the standards. But its magnitude is larger than the expected for the size of the firm.

The position of a firm is C: Although the magnitude of the feature we are interested in is near the expected for the size of the firm, the feature itself is below the standards.

The position of a firm is D: The feature itself is near the standards. However, its magnitude given the size of the firm is smaller than the expected.

The position of a firm is in between C and D: Both the feature and its magnitude, given the size of the firm, are below the expected. This is a frequent situation when assessing profitability. It means a firm too large for the generated earnings.

The position of a firm is E: The magnitude of the feature we are interested in is the expected for that size. But the feature itself is above the standards.

For example, the liquidity of a firm might be the expected one when considered as a contrast between short-term assets and liabilities. But both Current Assets and Liabilities might be smaller than the expected for the size of the firm. In that case, the rotated residual plot would show a position near "D" in figure 24.

A different way of reading positions in the rotated residual plot is also depicted in figure 24. It refers to quadrants, not to Cartesian axis. It leads to diagnostics based on log items, not on ratios. For the ratio y/x and s, an estimated size, we would have:

The firm lies in the first quadrant: Both x and y, the ratio components, are above their expected values for firms with that size.

The firm lies in the second quadrant: The denominator of the ratio, x, is below the expected and the numerator, y, is above the expected for firms of similar size.

The firm lies in the third quadrant: Both components are below the expected for a firm of that size. The firm is therefore oversized in what concerns those two items.

The firm lies in the fourth quadrant: The numerator, y, is above the expected for firms with that size. The denominator, x, is over-sized.

In the next section we give extensive examples of the use of rotated residual plots in financial diagnosis.

7.4 Using Complementary Ratios

We selected the Food Manufacturers industry and the profitability ratio to illustrate the joint use of complementary ratios and the possibilities they offer in the tracing of dynamic features. We shall use an abbreviation, RRP, to designate the rotated residual plot.

As we stressed in the last section, the RRP represents, in two dimensions, a ratio and its complement. Each zone of a RRP is assigned a financial diagnostic. It is the fact that a case lies in a particular zone that is important here. RRPs often exhibit irregular shapes. For example, RRPs reflecting profitability or flow of funds frequently show a comet-like shape, with a tail towards the third quadrant.

7.4.1 How to Build the RRP

For each year of the period 1983-1987 separately, we selected two samples. One contained firms with positive $EBIT$. The other one, those having negative $EBIT$. Then, a symmetric log transformation was applied to both — formula (1) on page 35 —. The items to be used, NW and $EBIT$, were then mean-adjusted year by year. Any trend was thus accounted for. The estimated size to be used, $\log s$, was also extracted as explained in section 6.1.

Next, we finded the Y-axis and the X-axis of the RRP, in the same way for the positive and negative $EBIT$ samples. Being

$$\varepsilon^{EBIT} = \log EBIT - \overline{\log EBIT} - \log s, \qquad \text{and} \qquad \varepsilon^{NW} = \log NW - \overline{\log NW} - \log s,$$

the two axis of the RRP are:

$$Y = \varepsilon^{EBIT} + \varepsilon^{NW} \qquad \text{and} \qquad X = \varepsilon^{EBIT} - \varepsilon^{NW}$$

As discussed in section 3.2, the positive-$EBIT$ and the negative-$EBIT$ scatters represent two different things and should be modelled separately. However, when studying dynamic features, it is convenient to introduce some form of continuity between them, thus allowing the tracing of firms whose $EBIT$ emerges from negative to positive values and conversely, during a period of several years.

Drawing the continuity between profits and losses: The third quadrant of the RRP containing positive-$EBIT$ firms seems to be logically linked with the first quadrant of the RRP for the negative ones. In figure 24 (page 79), the third quadrant is the region delimited by "C" and "D". It contains the firms showing a performance below the expected, both in profitability and in the magnitude of this feature when compared with the size of the firm. Similarly, in the samples showing negative $EBIT$, the first quadrant — the region between "E" and "B" in the same figure — contains firms with less severe losses and whose profitability is larger than the expected for negative-$EBIT$ firms of that size.

Let us suppose that a firm gradually falls into negative earnings. It draws in the RRP a trajectory towards the second or the third quadrant and then into the first or fourth quadrant of the negative-$EBIT$ plot. But the path through the third quadrant means a firm too large for what it is worth and also for the generated earnings. The path through the second quadrant would mean a firm too large for the generated earnings as well, but with a balanced capital regarding its size. Conversely, when a firm gradually emerges from a situation of poor profitability to a more healthy state, its path will go along the first or the fourth quadrant of the negative-$EBIT$ plot, until it reaches the third or the second one in the positive one. But the path through the first quadrant means an improvement in both earnings and capital regarding size.

The logical path linking these two plots, should link the worst situation of positive earnings with the best one of negative ones. The fact that firms often fall into negative earnings from quadrant other than the third one doesn't invalidate our reasoning. We also observed that the position of firms having, at least, one year with losses during the usual period, tends to be in the third quadrant. Given this, it seems as if the continuity between profits and losses should be drawn between the third quadrant of the positive-$EBIT$ plot and the first one of the negative one. Accordingly, we place the negative scatter in the third quadrant of the positive one, far away from the rest of the data. This is done by shifting all the cases in the former by a negative amount, and then mixing both plots, as shown in the RRP displayed in figure 25. Notice how the variability of the negative cases is much larger than the one of the positive ones.

7.4.2 Reading the RRP

In this section we compare the information conveyed by a set of ratios with the one of the RRP. The goal is to provide a means to get acquainted with this new tool by putting it side by side with the usual source of information for analysts, the ratios.

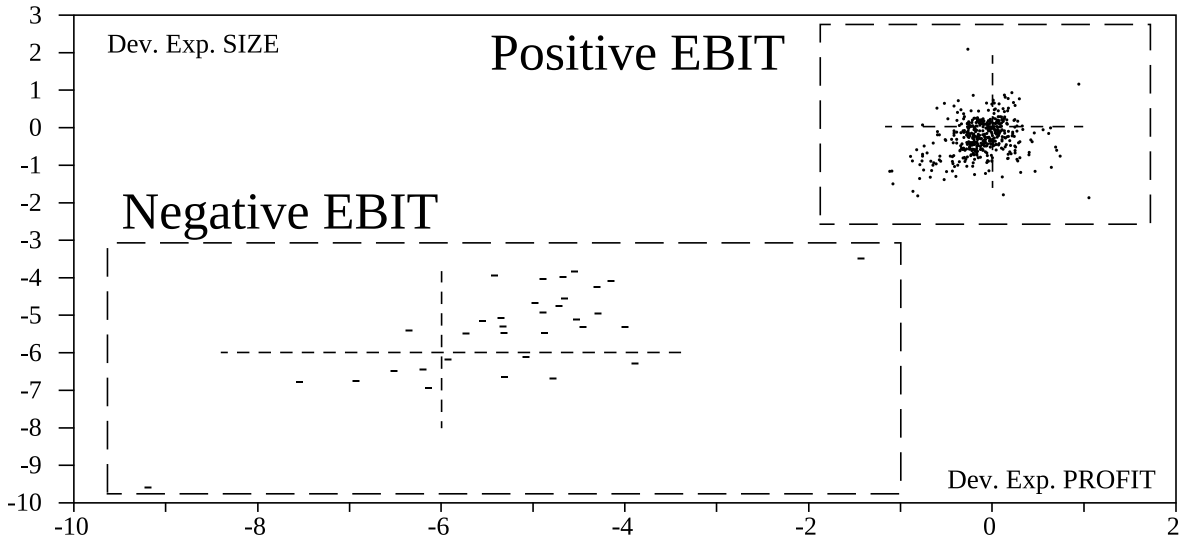

Figure 25: The joint RRP formed with the positive cases and the negative ones shifted so that they occupy a region of the third quadrant far away from the rest of the data.

We selected six firms from the Food industry. For each one of them we display the ratios Sales to Net Worth, Funds Flow to Total Debt, EBIT to Net Worth and also $\log s$, an estimated size. We also present the trajectories drawn in the RRP. The whole of the information is contained in figures 26 (page 84) and in figures 27 (page 85). This graphical description is complemented with table 35 in appendix (page 100). The figures show, on the left, a time-history of the mentioned ratios, and, on the right, the RRP. Each mark on the RRP refers to one year. For example, "4" shows the position of the firm in 1984. In order to interpret the RRP, notice that its axis measure percent deviations. The nearer a firm is from the centre, the less it diverges from the expected. The X-axis measures percent deviation from the profitability expected for the Food industry. The Y-axis measures how the magnitudes of Earnings and Net Worth, taken jointly, diverge from that expected for the size of the firm. In terms of quadrants, the first one means both Earnings and Net Worth larger than the expected. The second one means Net Worth larger than the expected but Earnings smaller, and so on.

We now comment on each one of the firms, organized in four classes: Firms occupying a steady position in the RRP, trends, other trajectories, and one case of negative EBIT.

Firms occupying a steady position: The firm on the top of figure 26 is an example of a steady position in the RRP. When reading the information conveyed by ratios about the evolution of UNITED BISCUITS, a large firm, we notice that during the period 1983-1987, the profitability ratios seem to suffer a small decrease. The reading of the RRP says that both the position of this firm in what concerns profitability, and the magnitude of this feature when compared with the size of the firm, are near the expected ones for the industry. They didn't change during the whole period.

Trends: NESTLE (UK), and MAUNDER (LLOYD), both in figure 26, show a trend towards a better performance during the five years considered. The first firm is small. It recovers from a position of profitability below the expected and over-sized regarding profits, to a new one agreeing with the standard for the industry. The second firm is larger than the expected for the industry. It improved its profitability from a standard position to above standards. Its Net Worth was kept near the expected

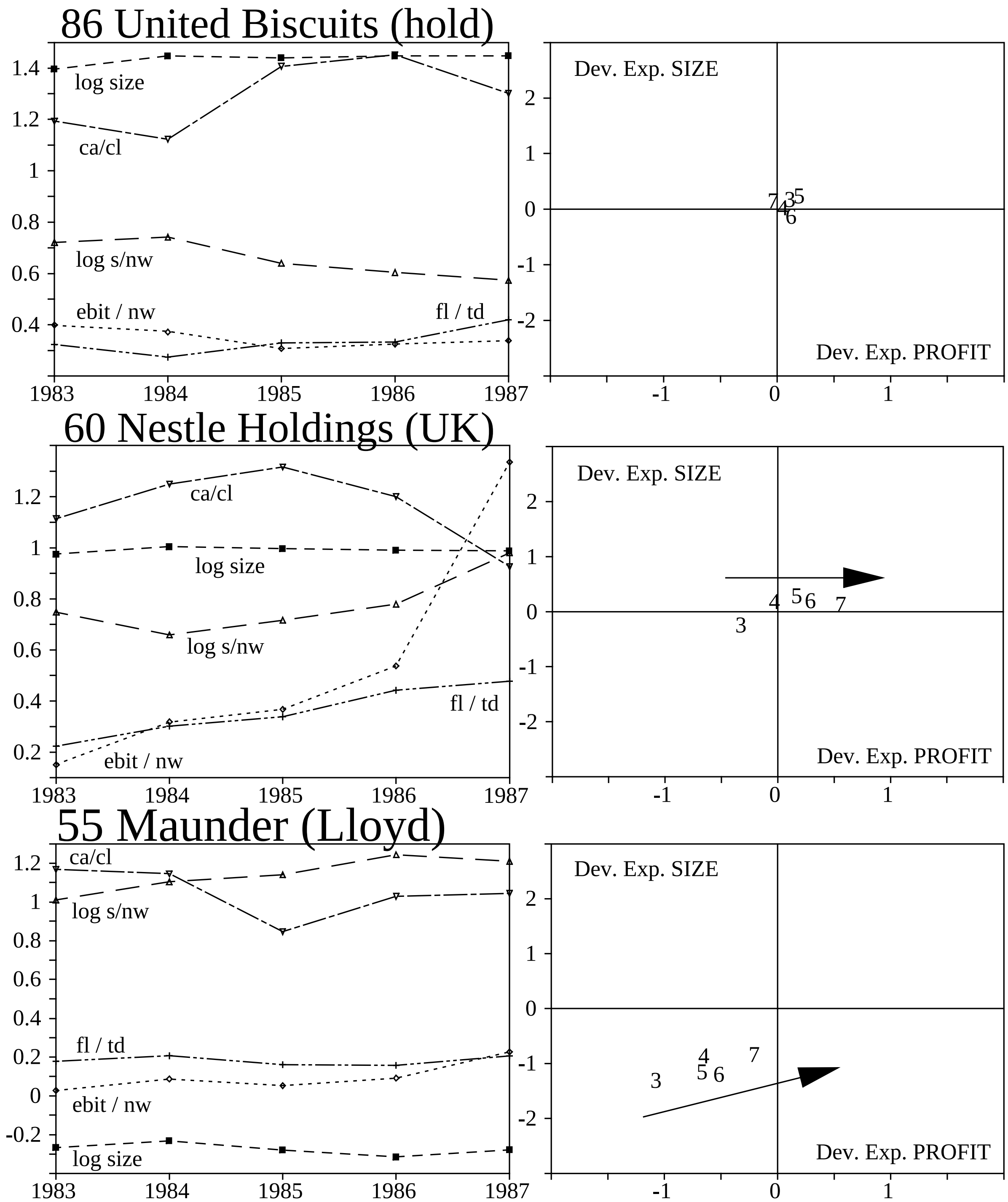

Figure 26: On the left, the evolution of a few ratios and a proxy for size during a period of five years. On the right, the corresponding rotated residual plot — EBIT to Net Worth — showing the trajectory drawn by the firm during the same period. Marks 1 to 7 indicate positions from 1983 to 1987.

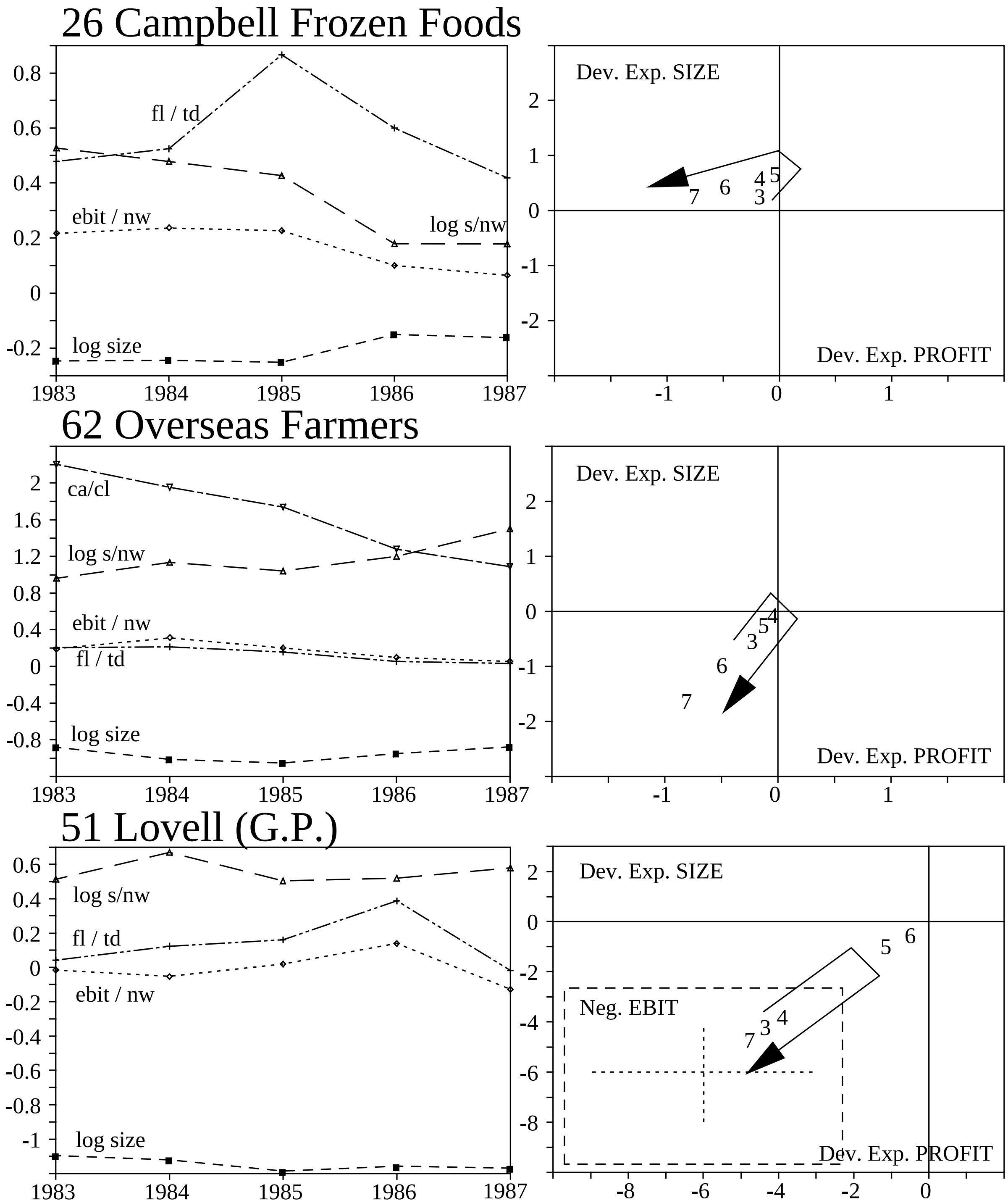

Figure 27: On the left, the evolution of a few ratios and a proxy for size during a period of five years. On the right, the corresponding rotated residual plot — EBIT to Net Worth — showing the trajectory drawn by the firm during the same period. Marks 1 to 7 indicate positions from 1983 to 1987.

for its size.

More complicated trends: CAMPBELL FROZEN FOODS (figure 27, page 85) is an example of a more complicated evolution. The whole of the trajectory lies in the upper two quadrants of the RRP. This means an excessive Net Worth when compared with the standards. The fact that this firm lies in the second quadrant, explicitly means that such a capital is not actually producing the expected profits.

OVERSEAS FARMERS is a typical case of increasingly poor profitability. Since the volume of sales didn't break down in the last two years of the period, other factors affected the performance of this firm. The RRP shows that this firm is too large for what it's worth, and also for the generated profits.

A case of negative $EBIT$: Figure 27 (page 85) also shows one firm having negative $EBIT$ except in 1985 and 1986. When compared with the standards for the industry, G. P. LOVELL, a small firm, is over-sized for what it's worth and for the profits it generates.

Discussion: When comparing the information conveyed by ratios with that of the RRP, it is clear that the former only tells part of a story. That is, the RRP is more specific in its diagnostic. For example, the profitability of NESTLE (UK) increased during the period. The RRP says as much, and also points out that such a gain was obtained purely by an increase in efficiency: When compared with the size of the firm, the Net Worth of NESTLE (UK) didn't improve. Also, ratios seem to suggest a small decrease in the profitability of UNITED BISCUITS during the period, whereas the RRP shows that, when taking size into account, this decrease is negligible. The degradation in the profitability of OVERSEAS FARMERS is explained by the RRP, more as a case of oversize regarding Net Worth, than as a lack of efficiency.

Finally, the RRP fully takes advantage of the widespread graphical power of computers, being suited for computerized analysis and machine learning to an extent ratios can't attain.

7.5 Summary

We have shown that the size-adjusted information left aside by the ratio y/x is the ratio $(y \times x)/s^2$, s being an estimated size. We also studied the potential interest of such a remainder in financial statement analysis. Then, we described bivariate tools incorporating relative size — or, in the case of the RRP, deviations from an expected size — and allowing the study of trajectories. The drawing of trajectories reveal a certain behaviour valuable for financial analysis and less explicit when using ratios solely. The RRP is a different, yet familiar, way of reading accounts. It is different from ratios in that it conveys two pieces of information at a time. But it is based on the same principles: A contrast between two magnitudes is supposed to capture features of the firm, and the value expected for the industry sets the standard.

Chapter 8

Conclusions and Tables

This study introduced a unified view that clarifies issues yet unsolved in financial statement analysis: The validity of ratios, the variety of their distributions, how to overcome their known limitations when used as input variables in statistical models.

Ratios are bivariate relations. Their distributions are determined by the ones of their components, plus the interaction between them. Thus, for understanding ratios, the first step consists of knowing the characteristics of such components. Raw data is much more regular than ratios. The observed items are two or three-parametric lognormal in cross-section. Lognormality is enough to explain the existence of outliers and the heteroscedasticity in models, often referred to in the literature. Regressions shouldn't be used to model lognormal variables, as large firms become strongly influential. Weighting is not an adequate remedy since it simply transfers the influence from the largest to the smallest cases in the sample. The trimming of outliers is also useless.

Items are the addition of two effects. The first one is common to a given report. It reflects the relative size of firms. The second one reflects a particular variability. Hence, items should be explained in terms of size and deviations from size.

Ratios can be extended so as to account for non-proportionality. Three-parametric lognormality demands the introduction of non-proportional terms in ratios. Such threshold ratios are natural extensions of the usual ones as both stem from the same stochastic mechanism. They seem promising for financial statement analysis: It is likely that they will be able to gather in one unique standard the features of firms having very different sizes.

Ratios are broadly lognormal. But accounting identities and other external forces can, in some cases, act as constraints, hiding the skewness of their distribution. This explains the major departures from lognormality observed in ratios. However, ratios are not exactly lognormal. In log space, the distribution of accounting data is determined by the interaction of a Gaussian common effect with the leptokurtic particular ones. In ratios, the particular ones are prevalent. In items, the Gaussian one dominates.

The size-adjusted information contained in two items can be expressed in terms of a ratio plus a remainder. Such a remainder is likely to be valuable for financial analysis but its building demands the use of an estimated size. Case-averages of selected items approach such an effect. This allows the assessment of the particular variability of items and the building of bivariate tools to be used in financial statement analysis.

Departures from two-parametric lognormality in five years	Items having only positive cases	Items having both + and −
Departures are never observed	69.2%	62.9%
A departure is observed once	11.0%	22.9%
A departure is observed twice	10.4%	11.4%
A departure is observed three times	5.5%	1.4%
A departure is observed four times	3.3%	1.4%
A departure is observed five times	0.5%	

In five years, deviations occurred: In item:	ONCE	TWICE	THREE TIMES	FOUR TIMES	FIVE TIMES
Sales	3		1	2	
Net Worth	2	3			
Wages	2	1			1
Inventory	3		1		
Debtors	1	3			
Creditors	2	1		1	
Current Assets	3	3			
Fixed Assets	1	2			
Total Assets		1	3		
Current Liabilities		1	3		
Number of employees	1	1	1	1	
Expenses	2		1	1	
Tot. Capital Empl.		3		1	
EBIT	2	2			
Operating Profit	2	2			
Long Term Debt	3	1		1	
Funds Flow Fr. Ops.	6	1	1		
Working Capital	3	2			

Table 22: Persistency of departures from the two-parameters model. By item.

In five years, deviations occurred: In industry:	ONCE	TWICE	THREE TIMES	FOUR TIMES	FIVE TIMES
Building Materials	3	1			
Metallurgy					
Paper and Packing	3	1	1		
Chemicals	4	1	1		
Electricity	3	3	2	1	
Industrial Plants	7	9	1	1	
Machine Tools	4				
Electronics	2	7	3	1	1
Motor Components	2				
Clothing	3	1			
Wool	2				
Textiles Mix.					
Leather					
Food Manufacturers	3	4	3	4	

Table 23: Persistency of departures from the two-parameters model. By industry.

		1983			1984			1985			1986			1987		
item	ind	N	W	P	N	W	P	N	W	P	N	W	P	N	W	P
S	BUIL	29	0.94	0.18	34	0.96	0.43	36	0.97	0.66	39	0.96	0.43	38	0.97	0.56
	METL	29	0.98	0.95	32	0.98	0.97	33	0.97	0.77	34	0.96	0.50	33	0.97	0.64
	PAPR	50	0.95	0.06	56	0.95	0.08	57	0.94	0.02	60	0.96	0.12	58	0.96	0.17
	CHEM	53	0.95	0.09	55	0.95	0.12	56	0.96	0.20	55	0.96	0.19	55	0.96	0.30
	ELEC	37	0.95	0.16	44	0.93	0.02	48	0.96	0.22	49	0.97	0.40	46	0.96	0.28
	I.PL	19	0.91	0.10	20	0.84	0.00	22	0.83	0.00	23	0.89	0.02	23	0.90	0.02
	TOOL	22	0.91	0.05	24	0.95	0.29	24	0.97	0.68	26	0.97	0.84	25	0.95	0.29
	ELTN	98	0.97	0.33	130	0.96	0.03	138	0.96	0.04	145	0.96	0.02	143	0.97	0.07
	MOTR	25	0.92	0.08	30	0.95	0.26	31	0.96	0.36	30	0.96	0.42	29	0.96	0.46
	CLOT	39	0.98	0.92	46	0.89	0	48	0.98	0.94	52	0.97	0.70	50	0.99	0.98
	WOOL	15	0.93	0.36	21	0.97	0.82	20	0.96	0.62	20	0.97	0.84	20	0.96	0.55
	TX.M	32	0.95	0.31	36	0.97	0.75	36	0.96	0.37	37	0.97	0.66	38	0.97	0.67
	LEAT	13	0.95	0.65	16	0.95	0.52	16	0.95	0.50	16	0.94	0.42	16	0.96	0.79
	FOOD	94	0.96	0.03	105	0.96	0.03	112	0.95	0.00	116	0.96	0.05	114	0.93	0
NW	BUIL	29	0.93	0.08	34	0.94	0.11	35	0.93	0.04	38	0.96	0.29	38	0.94	0.07
	METL	29	0.96	0.49	32	0.97	0.65	32	0.98	0.87	33	0.97	0.78	31	0.97	0.77
	PAPR	48	0.98	0.90	55	0.98	0.96	56	0.98	0.84	60	0.98	0.80	58	0.97	0.51
	CHEM	52	0.97	0.70	54	0.98	0.74	55	0.98	0.79	54	0.98	0.91	54	0.97	0.61
	ELEC	37	0.93	0.03	44	0.94	0.05	48	0.94	0.04	49	0.95	0.09	46	0.94	0.05
	I.PL	19	0.91	0.10	20	0.90	0.05	22	0.89	0.02	23	0.91	0.04	23	0.93	0.13
	TOOL	22	0.95	0.36	24	0.96	0.47	24	0.96	0.6	26	0.97	0.71	25	0.96	0.60
	ELTN	97	0.98	0.58	130	0.98	0.5	134	0.96	0.04	142	0.98	0.45	138	0.96	0.00
	MOTR	25	0.96	0.61	30	0.97	0.81	31	0.97	0.64	30	0.97	0.77	29	0.96	0.42
	CLOT	39	0.97	0.75	46	0.97	0.49	48	0.98	0.88	52	0.98	0.94	50	0.98	0.95
	WOOL	15	0.94	0.37	21	0.97	0.76	20	0.97	0.75	20	0.95	0.53	20	0.96	0.64
	TX.M	32	0.94	0.14	36	0.97	0.61	36	0.97	0.73	36	0.97	0.79	37	0.97	0.64
	LEAT	13	0.87	0.05	16	0.90	0.08	16	0.93	0.29	16	0.94	0.38	16	0.95	0.52
	FOOD	94	0.98	0.81	104	0.97	0.24	109	0.98	0.65	111	0.96	0.04	113	0.97	0.24
W	BUIL	29	0.97	0.72	34	0.97	0.64	36	0.97	0.64	39	0.97	0.53	38	0.97	0.57
	METL	29	0.97	0.65	32	0.95	0.26	33	0.95	0.23	34	0.96	0.49	33	0.95	0.25
	PAPR	50	0.97	0.58	55	0.97	0.38	56	0.97	0.42	59	0.96	0.23	57	0.96	0.31
	CHEM	52	0.95	0.09	54	0.96	0.16	55	0.96	0.20	54	0.96	0.33	54	0.97	0.40
	ELEC	37	0.94	0.10	44	0.94	0.05	48	0.94	0.04	49	0.95	0.06	46	0.94	0.05
	I.PL	19	0.94	0.32	20	0.89	0.04	22	0.93	0.12	23	0.93	0.12	23	0.95	0.45
	TOOL	22	0.93	0.15	24	0.93	0.14	24	0.93	0.10	26	0.94	0.16	25	0.94	0.18
	ELTN	97	0.95	0.02	130	0.96	0.00	138	0.93	0	144	0.94	0	143	0.94	0
	MOTR	25	0.95	0.33	30	0.95	0.28	31	0.95	0.33	30	0.95	0.34	29	0.96	0.43
	CLOT	39	0.97	0.75	45	0.99	0.99	48	0.97	0.66	52	0.97	0.45	50	0.98	0.73
	WOOL	15	0.94	0.47	21	0.98	0.93	20	0.97	0.89	20	0.98	0.96	20	0.97	0.88
	TX.M	31	0.94	0.16	35	0.96	0.29	35	0.95	0.15	36	0.94	0.09	37	0.95	0.18
	LEAT	13	0.95	0.68	16	0.91	0.15	16	0.92	0.22	16	0.93	0.26	16	0.93	0.26
	FOOD	92	0.97	0.40	105	0.96	0.02	111	0.96	0.02	116	0.96	0.05	114	0.96	0.08

Table 24: The Shapiro-Wilk test of lognormality. Items by industrial group and by year. N is the number of cases. First table.

		1983			1984			1985			1986			1987		
item	ind	N	W	P	N	W	P	N	W	P	N	W	P	N	W	P
I	BUIL	29	0.96	0.45	34	0.96	0.40	36	0.96	0.28	39	0.95	0.21	38	0.94	0.11
	METL	29	0.95	0.36	32	0.98	0.86	33	0.97	0.7	34	0.94	0.13	32	0.93	0.05
	PAPR	50	0.97	0.71	54	0.97	0.54	55	0.97	0.65	59	0.96	0.28	55	0.96	0.26
	CHEM	53	0.95	0.13	55	0.94	0.04	56	0.95	0.08	55	0.97	0.44	55	0.95	0.10
	ELEC	37	0.93	0.04	44	0.97	0.61	48	0.94	0.02	49	0.94	0.05	45	0.94	0.04
	I.PL	19	0.93	0.22	19	0.94	0.29	21	0.92	0.08	22	0.78	0	23	0.95	0.46
	TOOL	22	0.89	0.02	24	0.94	0.24	24	0.97	0.66	26	0.97	0.77	25	0.96	0.57
	ELTN	94	0.97	0.18	127	0.97	0.32	132	0.98	0.86	139	0.98	0.66	136	0.97	0.33
	MOTR	25	0.94	0.16	30	0.95	0.34	30	0.97	0.70	30	0.97	0.82	29	0.96	0.51
	CLOT	39	0.97	0.68	45	0.97	0.59	48	0.97	0.63	52	0.97	0.69	50	0.98	0.90
	WOOL	15	0.96	0.77	21	0.98	0.96	20	0.97	0.86	20	0.98	0.92	20	0.96	0.72
	TX.M	32	0.96	0.37	36	0.98	0.88	36	0.97	0.80	37	0.97	0.61	38	0.96	0.38
	LEAT	13	0.95	0.64	16	0.93	0.3	16	0.92	0.21	16	0.94	0.36	16	0.94	0.42
	FOOD	93	0.98	0.65	105	0.98	0.53	112	0.97	0.37	114	0.98	0.51	111	0.98	0.79
D	BUIL	29	0.96	0.52	34	0.98	0.82	36	0.98	0.81	39	0.98	0.93	38	0.98	0.90
	METL	29	0.96	0.47	32	0.95	0.21	33	0.97	0.67	34	0.97	0.55	33	0.97	0.65
	PAPR	50	0.95	0.07	55	0.94	0.03	56	0.94	0.02	59	0.95	0.07	56	0.95	0.06
	CHEM	53	0.96	0.19	55	0.96	0.14	56	0.95	0.10	55	0.96	0.14	55	0.96	0.18
	ELEC	37	0.95	0.14	44	0.97	0.58	48	0.96	0.20	49	0.97	0.45	46	0.95	0.11
	I.PL	19	0.97	0.78	20	0.89	0.02	22	0.88	0.01	23	0.94	0.2	23	0.92	0.07
	TOOL	22	0.94	0.30	24	0.89	0.01	24	0.95	0.38	26	0.96	0.43	25	0.97	0.68
	ELTN	98	0.98	0.70	130	0.97	0.31	138	0.96	0.06	145	0.96	0.04	143	0.96	0.03
	MOTR	25	0.95	0.37	30	0.95	0.21	31	0.94	0.17	30	0.97	0.59	29	0.96	0.52
	CLOT	39	0.97	0.71	46	0.97	0.57	48	0.97	0.56	52	0.97	0.41	50	0.98	0.97
	WOOL	15	0.93	0.30	21	0.97	0.79	20	0.97	0.89	20	0.98	0.96	20	0.94	0.36
	TX.M	32	0.96	0.52	36	0.98	0.91	36	0.97	0.68	37	0.97	0.59	38	0.97	0.79
	LEAT	13	0.90	0.14	16	0.92	0.18	16	0.92	0.23	16	0.92	0.21	16	0.91	0.12
	FOOD	93	0.97	0.46	104	0.98	0.62	112	0.97	0.43	116	0.97	0.44	113	0.97	0.12
C	BUIL	29	0.95	0.22	34	0.97	0.7	36	0.97	0.61	39	0.99	0.99	38	0.96	0.3
	METL	29	0.98	0.86	32	0.98	0.83	33	0.98	0.83	34	0.96	0.37	33	0.97	0.76
	PAPR	50	0.96	0.32	55	0.96	0.32	56	0.96	0.2	59	0.95	0.07	55	0.97	0.37
	CHEM	53	0.96	0.16	55	0.94	0.02	56	0.95	0.09	55	0.96	0.14	55	0.96	0.24
	ELEC	37	0.95	0.13	44	0.97	0.60	48	0.94	0.05	49	0.96	0.31	46	0.96	0.27
	I.PL	19	0.95	0.40	20	0.90	0.04	22	0.90	0.03	23	0.95	0.36	23	0.98	0.90
	TOOL	22	0.95	0.41	24	0.94	0.21	24	0.96	0.56	26	0.95	0.34	25	0.95	0.36
	ELTN	98	0.98	0.94	130	0.97	0.30	138	0.97	0.27	145	0.96	0.02	143	0.97	0.13
	MOTR	25	0.95	0.29	30	0.96	0.45	31	0.96	0.50	30	0.96	0.55	29	0.96	0.47
	CLOT	39	0.97	0.66	46	0.96	0.34	48	0.98	0.92	52	0.97	0.63	50	0.98	0.91
	WOOL	15	0.90	0.12	21	0.96	0.54	20	0.97	0.76	20	0.97	0.90	20	0.96	0.63
	TX.M	32	0.96	0.42	36	0.97	0.69	36	0.98	0.83	37	0.98	0.80	38	0.97	0.76
	LEAT	13	0.94	0.53	16	0.92	0.19	16	0.92	0.19	16	0.92	0.18	16	0.94	0.44
	FOOD	93	0.95	0.01	105	0.96	0.03	112	0.96	0.02	116	0.96	0.06	113	0.96	0.02

Table 25: The Shapiro-Wilk test of lognormality. Items by industrial group and by year. N is the number of cases. Second table.

		1983			1984			1985			1986			1987		
item	ind	N	W	P	N	W	P	N	W	P	N	W	P	N	W	P
CA	BUIL	29	0.95	0.36	34	0.97	0.62	36	0.96	0.41	38	0.96	0.33	38	0.95	0.13
	METL	29	0.97	0.67	32	0.97	0.75	33	0.97	0.68	34	0.96	0.46	33	0.96	0.52
	PAPR	50	0.96	0.30	54	0.95	0.11	57	0.83	0	60	0.97	0.35	58	0.96	0.29
	CHEM	52	0.96	0.18	55	0.95	0.05	56	0.95	0.07	54	0.97	0.36	55	0.95	0.07
	ELEC	37	0.94	0.07	44	0.94	0.05	48	0.93	0.01	49	0.94	0.05	46	0.94	0.02
	I.PL	19	0.91	0.09	20	0.90	0.05	22	0.89	0.02	23	0.92	0.06	23	0.96	0.53
	TOOL	21	0.90	0.03	24	0.93	0.13	24	0.96	0.47	26	0.98	0.88	25	0.97	0.80
	ELTN	97	0.98	0.82	127	0.98	0.71	133	0.96	0.05	142	0.96	0.00	141	0.96	0.01
	MOTR	25	0.94	0.19	30	0.96	0.39	31	0.97	0.58	30	0.97	0.67	29	0.97	0.69
	CLOT	39	0.98	0.83	46	0.97	0.68	48	0.98	0.77	52	0.98	0.71	50	0.98	0.73
	WOOL	15	0.95	0.65	21	0.98	0.93	20	0.98	0.93	20	0.97	0.82	20	0.96	0.71
	TX.M	32	0.93	0.08	36	0.96	0.40	36	0.97	0.68	37	0.97	0.73	37	0.97	0.51
	LEAT	13	0.94	0.49	16	0.94	0.47	16	0.94	0.34	16	0.94	0.41	16	0.94	0.37
	FOOD	93	0.96	0.11	105	0.97	0.13	112	0.94	0	115	0.97	0.20	112	0.96	0.02
FA	BUIL	29	0.92	0.04	34	0.95	0.28	36	0.95	0.15	39	0.93	0.04	38	0.94	0.11
	METL	29	0.96	0.50	32	0.97	0.74	33	0.96	0.37	34	0.97	0.53	32	0.96	0.42
	PAPR	50	0.97	0.37	56	0.97	0.61	57	0.98	0.95	60	0.98	0.94	57	0.98	0.80
	CHEM	53	0.95	0.09	55	0.95	0.09	56	0.95	0.08	55	0.94	0.03	55	0.95	0.09
	ELEC	37	0.95	0.17	44	0.95	0.15	48	0.95	0.16	49	0.96	0.19	46	0.94	0.05
	I.PL	19	0.93	0.24	20	0.90	0.06	22	0.90	0.03	22	0.90	0.04	23	0.93	0.17
	TOOL	22	0.95	0.43	24	0.95	0.38	24	0.95	0.40	26	0.94	0.18	25	0.95	0.36
	ELTN	98	0.98	0.84	130	0.98	0.54	138	0.98	0.78	145	0.97	0.35	143	0.97	0.08
	MOTR	25	0.95	0.32	30	0.98	0.89	30	0.97	0.73	30	0.98	0.86	29	0.97	0.78
	CLOT	39	0.97	0.74	46	0.96	0.31	48	0.97	0.45	52	0.96	0.27	50	0.97	0.41
	WOOL	15	0.95	0.61	21	0.97	0.90	20	0.98	0.99	20	0.99	0.99	20	0.98	0.93
	TX.M	32	0.96	0.53	36	0.96	0.34	36	0.95	0.26	36	0.96	0.36	37	0.96	0.32
	LEAT	13	0.90	0.13	16	0.93	0.28	16	0.96	0.67	16	0.95	0.51	16	0.96	0.72
	FOOD	94	0.98	0.69	105	0.97	0.32	112	0.97	0.24	116	0.97	0.45	114	0.97	0.23
TA	BUIL	29	0.94	0.12	34	0.96	0.39	36	0.95	0.24	39	0.94	0.07	38	0.94	0.06
	METL	29	0.97	0.76	32	0.98	0.88	33	0.97	0.74	34	0.97	0.71	33	0.97	0.55
	PAPR	50	0.96	0.36	56	0.96	0.15	57	0.96	0.24	60	0.96	0.10	58	0.97	0.46
	CHEM	53	0.96	0.16	55	0.95	0.09	56	0.95	0.10	55	0.96	0.21	55	0.96	0.19
	ELEC	37	0.93	0.04	44	0.94	0.06	48	0.93	0.02	49	0.95	0.09	46	0.94	0.04
	I.PL	19	0.91	0.10	20	0.91	0.08	22	0.90	0.03	23	0.90	0.03	23	0.96	0.59
	TOOL	22	0.93	0.15	24	0.94	0.25	24	0.97	0.78	26	0.96	0.45	25	0.96	0.59
	ELTN	98	0.99	0.99	130	0.97	0.26	137	0.96	0.02	144	0.96	0.00	142	0.95	0.00
	MOTR	25	0.94	0.21	30	0.97	0.61	31	0.97	0.74	30	0.96	0.44	29	0.95	0.33
	CLOT	39	0.97	0.74	46	0.98	0.93	48	0.96	0.31	52	0.98	0.81	50	0.98	0.83
	WOOL	15	0.94	0.4	21	0.98	0.93	20	0.98	0.96	20	0.98	0.92	20	0.97	0.77
	TX.M	32	0.93	0.08	36	0.96	0.47	36	0.97	0.51	37	0.96	0.46	37	0.96	0.42
	LEAT	13	0.91	0.19	16	0.92	0.17	16	0.93	0.29	16	0.94	0.36	16	0.96	0.75
	FOOD	94	0.96	0.05	105	0.96	0.03	112	0.94	0.00	116	0.96	0.10	113	0.96	0.03

Table 26: The Shapiro-Wilk test of lognormality. Items by industrial group and by year. N is the number of cases. Third table.

		1983			1984			1985			1986			1987		
item	ind	N	W	P	N	W	P	N	W	P	N	W	P	N	W	P
CL	BUIL	29	0.95	0.29	34	0.97	0.66	36	0.97	0.77	39	0.97	0.61	38	0.97	0.56
	METL	29	0.97	0.83	32	0.98	0.97	33	0.97	0.81	34	0.96	0.51	33	0.97	0.61
	PAPR	50	0.96	0.20	56	0.96	0.13	57	0.96	0.11	60	0.96	0.29	57	0.96	0.31
	CHEM	53	0.94	0.04	55	0.92	0.00	56	0.93	0.00	55	0.95	0.05	55	0.95	0.05
	ELEC	37	0.93	0.05	44	0.97	0.66	48	0.95	0.13	49	0.97	0.40	46	0.94	0.07
	I.PL	19	0.96	0.57	20	0.86	0.01	22	0.86	0.00	23	0.93	0.17	23	0.97	0.71
	TOOL	22	0.94	0.22	24	0.95	0.38	24	0.97	0.84	26	0.95	0.39	25	0.96	0.64
	ELTN	98	0.97	0.35	130	0.97	0.21	138	0.96	0.02	145	0.95	0	143	0.95	0.00
	MOTR	25	0.95	0.42	30	0.95	0.30	31	0.97	0.62	30	0.95	0.27	29	0.96	0.48
	CLOT	39	0.98	0.88	46	0.97	0.72	48	0.97	0.57	52	0.97	0.61	50	0.98	0.96
	WOOL	15	0.90	0.09	21	0.96	0.70	20	0.98	0.97	20	0.97	0.83	20	0.96	0.66
	TX.M	32	0.97	0.56	36	0.97	0.8	36	0.97	0.79	37	0.98	0.93	38	0.98	0.92
	LEAT	13	0.95	0.67	16	0.94	0.43	16	0.92	0.23	16	0.94	0.41	16	0.96	0.64
	FOOD	94	0.95	0.01	105	0.96	0.02	112	0.95	0.01	116	0.96	0.06	113	0.96	0.05
N	BUIL	28	0.97	0.64	34	0.96	0.5	36	0.97	0.61	39	0.97	0.50	38	0.96	0.44
	METL	29	0.97	0.73	32	0.95	0.19	33	0.94	0.13	34	0.95	0.20	33	0.94	0.15
	PAPR	50	0.96	0.35	55	0.98	0.82	56	0.97	0.62	59	0.97	0.52	57	0.97	0.59
	CHEM	52	0.97	0.70	54	0.97	0.36	55	0.97	0.46	54	0.97	0.61	53	0.97	0.53
	ELEC	36	0.94	0.11	43	0.96	0.20	47	0.95	0.10	49	0.94	0.03	46	0.94	0.02
	I.PL	19	0.92	0.14	20	0.87	0.01	22	0.90	0.03	23	0.91	0.04	23	0.94	0.18
	TOOL	22	0.91	0.07	24	0.94	0.19	24	0.92	0.06	26	0.91	0.03	25	0.92	0.05
	ELTN	97	0.96	0.10	130	0.95	0.00	138	0.94	0	145	0.94	0	143	0.94	0
	MOTR	25	0.95	0.42	30	0.95	0.31	30	0.96	0.53	30	0.96	0.52	29	0.96	0.55
	CLOT	39	0.98	0.88	45	0.99	0.99	48	0.98	0.77	52	0.98	0.78	50	0.98	0.86
	WOOL	15	0.95	0.57	21	0.98	0.99	20	0.99	0.99	20	0.99	0.99	20	0.98	0.99
	TX.M	32	0.95	0.19	36	0.97	0.73	36	0.96	0.45	37	0.96	0.31	38	0.97	0.56
	LEAT	13	0.92	0.29	16	0.91	0.12	16	0.91	0.16	16	0.92	0.17	16	0.92	0.21
	FOOD	92	0.97	0.50	105	0.97	0.26	111	0.97	0.29	116	0.98	0.53	113	0.98	0.50
EX	BUIL	29	0.93	0.08	34	0.96	0.36	36	0.97	0.58	39	0.97	0.60	38	0.97	0.77
	METL	29	0.98	0.98	32	0.98	0.97	33	0.95	0.23	34	0.96	0.47	33	0.97	0.80
	PAPR	49	0.96	0.18	56	0.94	0.01	57	0.90	0	60	0.95	0.06	58	0.94	0.02
	CHEM	53	0.96	0.20	55	0.96	0.21	56	0.96	0.14	55	0.96	0.25	55	0.97	0.44
	ELEC	37	0.95	0.23	44	0.97	0.65	47	0.96	0.27	49	0.96	0.35	46	0.96	0.29
	I.PL	19	0.92	0.13	20	0.87	0.01	22	0.91	0.05	23	0.92	0.10	23	0.93	0.12
	TOOL	22	0.93	0.14	24	0.95	0.32	24	0.97	0.74	26	0.98	0.92	25	0.95	0.33
	ELTN	98	0.98	0.84	130	0.98	0.75	138	0.98	0.80	145	0.98	0.67	143	0.98	0.87
	MOTR	25	0.92	0.07	30	0.95	0.28	31	0.95	0.22	30	0.96	0.43	29	0.95	0.25
	CLOT	39	0.98	0.86	46	0.84	0	48	0.97	0.69	52	0.97	0.69	50	0.98	0.97
	WOOL	15	0.96	0.70	21	0.98	0.95	20	0.98	0.92	20	0.97	0.88	20	0.97	0.78
	TX.M	32	0.97	0.80	36	0.98	0.87	36	0.96	0.49	37	0.98	0.92	38	0.98	0.84
	LEAT	13	0.95	0.63	16	0.96	0.73	16	0.95	0.54	16	0.93	0.31	16	0.95	0.61
	FOOD	94	0.96	0.06	105	0.96	0.02	112	0.95	0.00	116	0.96	0.04	114	0.94	0

Table 27: The Shapiro-Wilk test of lognormality. Items by industrial group and by year. N is the number of cases. Fourth table.

		1983			1984			1985			1986			1987		
item	ind	N	W	P	N	W	P	N	W	P	N	W	P	N	W	P
TC	BUIL	29	0.94	0.15	34	0.94	0.12	35	0.93	0.06	38	0.95	0.13	38	0.94	0.08
	METL	29	0.96	0.47	32	0.96	0.34	33	0.98	0.88	34	0.98	0.86	32	0.98	0.92
	PAPR	49	0.98	0.91	55	0.98	0.74	57	0.97	0.63	60	0.97	0.52	58	0.97	0.61
	CHEM	53	0.97	0.57	55	0.97	0.46	56	0.97	0.38	55	0.97	0.62	55	0.97	0.35
	ELEC	37	0.93	0.02	44	0.94	0.04	48	0.94	0.04	49	0.95	0.06	46	0.94	0.04
	I.PL	19	0.90	0.07	20	0.90	0.06	22	0.88	0.01	23	0.90	0.03	23	0.92	0.10
	TOOL	22	0.93	0.19	24	0.95	0.28	24	0.95	0.40	26	0.97	0.72	25	0.95	0.41
	ELTN	97	0.98	0.72	130	0.98	0.62	135	0.96	0.01	143	0.97	0.10	138	0.95	0
	MOTR	25	0.95	0.30	30	0.97	0.81	31	0.97	0.75	30	0.96	0.51	29	0.95	0.24
	CLOT	39	0.98	0.93	46	0.97	0.42	48	0.98	0.88	52	0.98	0.94	50	0.98	0.77
	WOOL	15	0.95	0.63	21	0.96	0.66	20	0.96	0.69	20	0.95	0.53	20	0.95	0.50
	TX.M	32	0.93	0.09	36	0.97	0.58	36	0.97	0.61	36	0.97	0.59	37	0.97	0.57
	LEAT	13	0.88	0.09	16	0.90	0.11	16	0.93	0.26	16	0.94	0.41	16	0.95	0.60
	FOOD	94	0.98	0.56	105	0.97	0.48	111	0.96	0.04	112	0.96	0.02	113	0.96	0.05
EB	BUIL	29	0.96	0.43	33	0.88	0.00	36	0.97	0.59	38	0.97	0.49	38	0.94	0.07
	METL	25	0.96	0.56	29	0.97	0.82	27	0.95	0.28	31	0.98	0.84	29	0.95	0.30
	PAPR	45	0.98	0.84	53	0.99	0.99	57	0.97	0.50	59	0.97	0.46	57	0.97	0.56
	CHEM	51	0.96	0.18	54	0.97	0.47	54	0.98	0.80	53	0.97	0.69	52	0.95	0.11
	ELEC	37	0.98	0.84	41	0.98	0.91	47	0.96	0.26	46	0.98	0.92	43	0.96	0.33
	I.PL	17	0.85	0.01	18	0.90	0.05	21	0.92	0.12	19	0.94	0.27	19	0.92	0.13
	TOOL	17	0.91	0.14	23	0.98	0.95	22	0.97	0.72	24	0.97	0.71	25	0.96	0.59
	ELTN	91	0.98	0.76	122	0.97	0.28	121	0.96	0.05	121	0.95	0.00	128	0.95	0.00
	MOTR	22	0.95	0.38	28	0.98	0.87	27	0.97	0.67	27	0.96	0.55	27	0.97	0.80
	CLOT	36	0.96	0.41	39	0.97	0.69	44	0.97	0.50	50	0.98	0.90	45	0.98	0.93
	WOOL	15	0.95	0.57	21	0.96	0.69	20	0.96	0.60	20	0.97	0.88	20	0.98	0.97
	TX.M	29	0.93	0.07	33	0.96	0.45	32	0.95	0.26	34	0.95	0.26	36	0.97	0.79
	LEAT	13	0.91	0.20	15	0.95	0.60	15	0.96	0.73	16	0.95	0.62	16	0.95	0.62
	FOOD	87	0.96	0.10	97	0.96	0.02	106	0.97	0.16	107	0.97	0.12	106	0.93	0
OP	BUIL	29	0.96	0.53	33	0.96	0.54	34	0.97	0.68	36	0.97	0.75	37	0.94	0.06
	METL	24	0.95	0.38	28	0.96	0.50	26	0.97	0.73	31	0.98	0.89	28	0.97	0.67
	PAPR	44	0.98	0.86	51	0.95	0.14	55	0.96	0.21	58	0.97	0.55	57	0.97	0.43
	CHEM	49	0.97	0.58	54	0.97	0.59	53	0.97	0.65	51	0.97	0.63	52	0.97	0.37
	ELEC	33	0.97	0.79	40	0.98	0.80	45	0.97	0.49	43	0.97	0.58	41	0.96	0.39
	I.PL	17	0.86	0.02	18	0.88	0.03	19	0.92	0.14	18	0.91	0.11	16	0.94	0.46
	TOOL	16	0.95	0.60	21	0.97	0.85	20	0.97	0.86	23	0.97	0.70	24	0.98	0.90
	ELTN	87	0.98	0.81	118	0.97	0.47	121	0.96	0.09	120	0.96	0.04	126	0.96	0.01
	MOTR	22	0.95	0.36	26	0.97	0.65	27	0.96	0.61	25	0.97	0.75	26	0.96	0.47
	CLOT	35	0.95	0.14	38	0.97	0.51	44	0.86	0	50	0.97	0.70	43	0.98	0.8
	WOOL	14	0.96	0.72	21	0.96	0.65	20	0.97	0.87	20	0.98	0.95	20	0.96	0.56
	TX.M	27	0.95	0.25	31	0.96	0.50	32	0.96	0.55	32	0.96	0.46	35	0.98	0.93
	LEAT	13	0.91	0.18	15	0.95	0.57	15	0.95	0.64	16	0.97	0.86	16	0.96	0.78
	FOOD	87	0.98	0.64	95	0.96	0.05	104	0.97	0.26	104	0.97	0.27	98	0.95	0.00

Table 28: The Shapiro-Wilk test of lognormality. Items by industrial group and by year. N is the number of cases. Fifth table.

		1983			1984			1985			1986			1987		
item	ind	N	W	P	N	W	P	N	W	P	N	W	P	N	W	P
DB	BUIL	22	0.97	0.78	28	0.97	0.75	29	0.97	0.83	34	0.94	0.13	34	0.95	0.23
	METL	19	0.96	0.66	21	0.97	0.77	20	0.97	0.82	21	0.93	0.15	19	0.97	0.80
	PAPR	37	0.97	0.75	43	0.96	0.29	42	0.98	0.80	46	0.97	0.56	47	0.95	0.17
	CHEM	36	0.94	0.11	39	0.96	0.42	42	0.96	0.26	43	0.93	0.01	43	0.93	0.02
	ELEC	19	0.93	0.25	27	0.95	0.23	29	0.95	0.35	33	0.97	0.57	33	0.97	0.63
	I.PL	12	0.97	0.87	14	0.85	0.02	16	0.90	0.11	17	0.93	0.31	18	0.95	0.52
	TOOL	15	0.88	0.05	19	0.96	0.60	20	0.96	0.60	21	0.96	0.57	21	0.91	0.06
	ELTN	55	0.97	0.42	78	0.97	0.22	94	0.98	0.82	101	0.97	0.22	100	0.98	0.55
	MOTR	21	0.89	0.02	26	0.95	0.29	25	0.95	0.35	28	0.94	0.21	27	0.96	0.43
	CLOT	25	0.94	0.22	29	0.93	0.07	33	0.94	0.13	37	0.95	0.25	35	0.98	0.86
	WOOL	7	0.78	0.03	10	0.95	0.72	11	0.95	0.72	12	0.97	0.87	12	0.97	0.93
	TX.M	20	0.96	0.56	23	0.97	0.77	23	0.95	0.40	26	0.96	0.52	26	0.95	0.24
	LEAT	6	0.93	0.58	8	0.97	0.91	9	0.90	0.28	9	0.95	0.77	11	0.94	0.57
	FOOD	64	0.95	0.04	74	0.95	0.03	86	0.95	0.01	90	0.93	0	84	0.97	0.26
FL	BUIL	29	0.96	0.43	34	0.91	0.01	36	0.96	0.27	39	0.95	0.23	38	0.95	0.13
	METL	28	0.96	0.59	31	0.97	0.80	31	0.98	0.93	31	0.98	0.94	29	0.96	0.37
	PAPR	48	0.97	0.70	56	0.98	0.92	57	0.97	0.36	60	0.97	0.45	57	0.97	0.44
	CHEM	51	0.95	0.09	55	0.96	0.21	54	0.96	0.16	53	0.96	0.19	54	0.94	0.03
	ELEC	37	0.96	0.37	42	0.98	0.78	48	0.95	0.10	48	0.93	0.02	45	0.95	0.11
	I.PL	17	0.86	0.02	18	0.90	0.06	21	0.91	0.06	19	0.93	0.23	20	0.98	0.98
	TOOL	19	0.95	0.52	24	0.96	0.65	24	0.99	0.99	25	0.94	0.25	24	0.96	0.50
	ELTN	91	0.97	0.23	123	0.97	0.12	126	0.96	0.07	130	0.97	0.22	130	0.95	0.00
	MOTR	22	0.96	0.60	30	0.97	0.82	29	0.97	0.64	28	0.96	0.40	28	0.92	0.04
	CLOT	36	0.97	0.78	40	0.93	0.04	44	0.99	0.98	50	0.97	0.68	45	0.89	0
	WOOL	15	0.94	0.44	21	0.96	0.65	20	0.95	0.53	20	0.96	0.74	20	0.98	0.94
	TX.M	29	0.94	0.12	34	0.96	0.38	33	0.95	0.27	34	0.95	0.17	36	0.97	0.63
	LEAT	13	0.92	0.25	16	0.97	0.93	15	0.95	0.54	16	0.96	0.73	16	0.96	0.70
	FOOD	92	0.97	0.52	101	0.96	0.03	109	0.96	0.03	113	0.97	0.25	108	0.95	0.00
WC	BUIL	28	0.95	0.24	30	0.98	0.89	32	0.96	0.36	37	0.97	0.62	35	0.97	0.75
	METL	27	0.96	0.54	30	0.94	0.17	32	0.97	0.60	33	0.96	0.34	31	0.95	0.19
	PAPR	43	0.93	0.01	48	0.97	0.51	50	0.96	0.16	54	0.97	0.51	54	0.97	0.41
	CHEM	47	0.98	0.90	50	0.98	0.83	53	0.98	0.75	49	0.95	0.10	50	0.97	0.40
	ELEC	36	0.97	0.55	42	0.97	0.57	47	0.98	0.85	46	0.98	0.97	42	0.95	0.12
	I.PL	19	0.89	0.03	19	0.90	0.07	21	0.89	0.02	23	0.94	0.23	21	0.93	0.20
	TOOL	21	0.95	0.49	24	0.95	0.28	24	0.94	0.27	26	0.97	0.68	24	0.96	0.61
	ELTN	89	0.98	0.56	117	0.96	0.09	113	0.96	0.02	129	0.97	0.35	127	0.95	0.00
	MOTR	25	0.95	0.39	29	0.96	0.49	30	0.96	0.49	27	0.95	0.32	28	0.97	0.62
	CLOT	37	0.96	0.27	45	0.96	0.21	44	0.98	0.86	51	0.97	0.62	49	0.98	0.84
	WOOL	13	0.95	0.64	19	0.95	0.40	18	0.94	0.35	18	0.91	0.11	18	0.89	0.04
	TX.M	29	0.92	0.05	34	0.95	0.26	33	0.95	0.25	34	0.94	0.08	35	0.95	0.17
	LEAT	12	0.96	0.73	15	0.95	0.60	15	0.95	0.58	14	0.97	0.84	15	0.89	0.06
	FOOD	79	0.97	0.54	85	0.96	0.12	98	0.98	0.84	100	0.98	0.68	97	0.95	0.00

Table 29: The Shapiro-Wilk test of lognormality. Items by industrial group and by year. N is the number of cases. Sixth table.

1983: In 1983 there were no bad cases.

1984: There are bad cases in two industries:

- **CLOTHING:** Sales and Expenses. The firm STORMGARD PLC sold 54 (units are thousands of pounds) and had also very small expenses. The sample has only 46 cases and the next smallest value of sales is 2402. STORMGARD turns out to be a strong outlier in a very small sample. We further notice that this firm has sales which are larger than earnings.
- **ELECTRONICS:** Wages and Number of Employees. There are three clear clusters. A cluster of eight large firms is clearly detached from the rest of the distribution. It is interesting to notice that neither the skewness nor the kurtosis exhibit values far from the acceptable.

1985: There are bad cases in three industries:

- **PAPER AND PACKING:** Current Assets and Expenses. In 57 cases there is one firm, EAST LANCASHIRE PAPER GRO, with $CA = 1$. The next smallest value in the sample is 1265. The same as Expenses. EAST LANCASHIRE is also one of the 5 firms exhibiting EBIT larger than Sales during one or two years of the period.
- **FOOD:** Current Assets. There are three clear clusters. Again, Skewness and Kurtosis are unable to trace the irregular shape of this distribution.
- **CLOTHING:** Operating Profit. The firm UNIGROUP PLC appears in the database with $OPP = 1$, and such a profit turns out to be a strong outlier in a small group. The next smallest case is $OPP = 52$. The sample has 44 cases.

1986: There are bad cases in three industries:

- **ELECTRONICS:** Wages, Number of Employees and Current Liabilities. Again three clusters, large firms well separated from the distribution. Skewness and kurtosis are normal.
- **INDUSTRIAL PLANTS:** Inventory. BIMEC PLC has $I = 1$. Next smallest value, 278. In a sample of 22 cases this is enough to influence normality tests.
- **FOOD:** Long term Debt. Very clear three-modal distribution.

1987: There are bad cases in three industries:

- **FOOD:** Sales, Expenses and Earnings. Again, three very clear clusters but in this case the cluster of small firms is detached from the others. Then, there is also a peaking central cluster. Skewness and kurtosis again fail to trace the lack of normality. In this group there are four firms with EBIT larger than Sales.
- **ELECTRONICS:** Wages, Number of Employees and Total Capital Employed. There are three clusters as in two previous years. But the group of eight very large firms is now less detached from the others than it was in previous years.
- **CLOTHING:** Funds Flow From Operations. The firm GOODMAN GROUP PLC displays a $FL = 9$ which is a clear outlier. The next smallest cases have $FL = 339$. The sample has 45 cases.

Figure 28: Description of all the extreme departures from lognormality referred to in the main text as the “bad samples”.

item	ind	1983	1984	1985	1986	1987	item	ind	1983	1984	1985	1986	1987
C	1	0.89	0.9	0.91	0.92	0.91	FA	1	0.9	0.89	0.85	0.86	0.88
	2	0.8	0.86	0.81	0.74	0.77		2	0.62	0.5	0.6	0.63	0.6
	3	0.89	0.89	0.92	0.93	0.94		3	0.76	0.76	0.72	0.83	0.83
	4	0.91	0.92	0.9	0.9	0.88		4	0.89	0.86	0.86	0.81	0.8
	5	0.95	0.96	0.96	0.95	0.94		5	0.81	0.72	0.74	0.74	0.71
	6	0.86	0.91	0.93	0.95	0.87		6	0.82	0.81	0.83	0.87	0.8
	7	0.92	0.96	0.92	0.9	0.91		7	0.82	0.83	0.85	0.88	0.81
	8	0.9	0.9	0.9	0.9	0.91		8	0.75	0.83	0.85	0.84	0.81
	9	0.94	0.86	0.93	0.92	0.92		9	0.94	0.92	0.96	0.93	0.91
	10	0.88	0.9	0.88	0.84	0.8		10	0.76	0.82	0.75	0.73	0.76
	11	0.8	0.87	0.86	0.85	0.89		11	0.67	0.82	0.83	0.87	0.87
	12	0.91	0.93	0.96	0.95	0.94		12	0.78	0.81	0.83	0.85	0.83
	13	0.92	0.94	0.96	0.95	0.95		13	0.82	0.88	0.9	0.87	0.88
	14	0.93	0.94	0.93	0.94	0.94		14	0.84	0.9	0.89	0.87	0.84
D	1	0.96	0.97	0.96	0.97	0.97	FL	1	0.86	0.79	0.9	0.87	0.9
	2	0.85	0.89	0.84	0.79	0.78		2	0.73	0.81	0.81	0.83	0.75
	3	0.9	0.84	0.83	0.87	0.87		3	0.85	0.82	0.9	0.83	0.91
	4	0.9	0.89	0.91	0.85	0.87		4	0.91	0.89	0.9	0.88	0.67
	5	0.93	0.92	0.91	0.92	0.92		5	0.88	0.88	0.86	0.82	0.8
	6	0.92	0.95	0.95	0.95	0.89		6	0.81	0.82	0.82	0.84	0.6
	7	0.86	0.88	0.87	0.83	0.71		7	0.52	0.75	0.71	0.58	0.91
	8	0.91	0.92	0.95	0.95	0.95		8	0.87	0.89	0.84	0.82	0.91
	9	0.98	0.97	0.93	0.97	0.96		9	0.95	0.84	0.87	0.91	0.85
	10	0.85	0.84	0.78	0.72	0.78		10	0.76	0.6	0.78	0.6	0.63
	11	0.85	0.89	0.91	0.91	0.9		11	0.78	0.88	0.89	0.87	0.89
	12	0.93	0.95	0.95	0.95	0.94		12	0.92	0.87	0.95	0.94	0.94
	13	0.8	0.86	0.89	0.88	0.87		13	0.96	0.9	0.96	0.96	0.96
	14	0.95	0.93	0.94	0.94	0.94		14	0.89	0.92	0.89	0.91	0.91
I	1	0.93	0.94	0.92	0.92	0.92	NW	1	0.85	0.9	0.89	0.9	0.9
	2	0.84	0.87	0.79	0.87	0.69		2	0.7	0.82	0.8	0.8	0.77
	3	0.95	0.95	0.94	0.79	0.8		3	0.89	0.81	0.79	0.8	0.89
	4	0.82	0.8	0.79	0.83	0.78		4	0.88	0.9	0.9	0.89	0.88
	5	0.93	0.91	0.89	0.91	0.91		5	0.9	0.85	0.85	0.84	0.83
	6	0.83	0.85	0.91	0.91	0.9		6	0.75	0.77	0.84	0.86	0.79
	7	0.87	0.88	0.87	0.9	0.92		7	0.84	0.75	0.71	0.82	0.82
	8	0.82	0.81	0.78	0.82	0.73		8	0.94	0.91	0.93	0.93	0.92
	9	0.98	0.92	0.94	0.94	0.92		9	0.94	0.95	0.91	0.92	0.9
	10	0.91	0.9	0.88	0.87	0.89		10	0.81	0.84	0.84	0.85	0.87
	11	0.75	0.84	0.78	0.8	0.73		11	0.77	0.85	0.89	0.89	0.89
	12	0.94	0.96	0.95	0.93	0.93		12	0.91	0.93	0.94	0.94	0.94
	13	0.95	0.98	0.98	0.98	0.97		13	0.93	0.96	0.97	0.97	0.96
	14	0.91	0.93	0.92	0.88	0.87		14	0.89	0.88	0.88	0.92	0.88

Table 30: Proportion of explained variability, R^2, when an estimated size explains twelve accounting items by industry and by year. First Table. Numbers from 1 to 14 identify industries: 1, Building Materials; 2, Metallurgy; 3, Paper and Packing; 4, Chemicals; 5, Electrical; 6, Industrial Plants; 7, Machine Tools; 8, Electronics; 9, Motor Components; 10, Clothing; 11, Wool, 12, Miscellaneous Textiles, 13, Leather; 14, Food Manufacturers.

item	ind	1983	1984	1985	1986	1987	item	ind	1983	1984	1985	1986	1987
S	1	0.97	0.98	0.98	0.98	0.98	WC	1	0.64	0.81	0.75	0.78	0.75
	2	0.64	0.66	0.65	0.68	0.64		2	0.87	0.87	0.77	0.75	0.82
	3	0.95	0.88	0.94	0.9	0.91		3	0.52	0.75	0.74	0.77	0.81
	4	0.9	0.88	0.88	0.86	0.86		4	0.87	0.83	0.81	0.61	0.71
	5	0.96	0.9	0.94	0.93	0.94		5	0.61	0.81	0.74	0.62	0.82
	6	0.97	0.98	0.96	0.97	0.94		6	0.68	0.69	0.7	0.56	0.76
	7	0.94	0.97	0.96	0.95	0.92		7	0.72	0.7	0.64	0.84	0.78
	8	0.94	0.94	0.93	0.93	0.93		8	0.85	0.87	0.88	0.8	0.84
	9	0.99	0.95	0.97	0.96	0.95		9	0.75	0.86	0.86	0.78	0.79
	10	0.96	0.9	0.96	0.91	0.92		10	0.72	0.49	0.62	0.56	0.41
	11	0.86	0.93	0.93	0.92	0.94		11	0.68	0.84	0.83	0.83	0.73
	12	0.95	0.96	0.97	0.97	0.97		12	0.88	0.79	0.58	0.87	0.81
	13	0.99	0.99	0.98	0.98	0.98		13	0.91	0.96	0.92	0.92	0.82
	14	0.91	0.93	0.92	0.92	0.89		14	0.8	0.8	0.74	0.78	0.74
W	1	0.98	0.98	0.98	0.98	0.97	DEBT	1	0.55	0.68	0.69	0.55	0.65
	2	0.77	0.78	0.83	0.85	0.86		2	0.38	0.31	0.38	0.43	0.5
	3	0.94	0.93	0.92	0.92	0.93		3	0.51	0.27	0.4	0.5	0.63
	4	0.95	0.95	0.95	0.96	0.95		4	0.73	0.73	0.63	0.52	0.53
	5	0.95	0.92	0.88	0.95	0.94		5	0.73	0.49	0.46	0.41	0.36
	6	0.96	0.98	0.96	0.97	0.94		6	0.23	0.86	0.86	0.6	0.33
	7	0.93	0.97	0.96	0.96	0.94		7	0.04	0.34	0.24	0.37	0.58
	8	0.88	0.9	0.91	0.92	0.91		8	0.54	0.62	0.55	0.49	0.49
	9	0.98	0.95	0.97	0.95	0.97		9	0.72	0.92	0.83	0.68	0.66
	10	0.88	0.88	0.87	0.88	0.88		10	0.34	0.26	0.39	0.32	0.29
	11	0.77	0.87	0.89	0.91	0.9		11	0.12	0.31	0.54	0.52	0.59
	12	0.88	0.9	0.92	0.93	0.92		12	0.71	0.74	0.83	0.84	0.76
	13	0.93	0.95	0.94	0.93	0.91		13	0.65	0.85	0.72	0.66	0.54
	14	0.93	0.94	0.95	0.95	0.95		14	0.82	0.77	0.73	0.68	0.66
CA	1	0.98	0.99	0.95	0.98	0.99	EBIT	1	0.82	0.78	0.82	0.82	0.88
	2	0.88	0.91	0.9	0.91	0.93		2	0.75	0.83	0.78	0.83	0.84
	3	0.95	0.97	0.96	0.96	0.97		3	0.89	0.77	0.83	0.86	0.88
	4	0.97	0.93	0.94	0.91	0.87		4	0.86	0.89	0.83	0.85	0.81
	5	0.98	0.96	0.97	0.97	0.97		5	0.75	0.82	0.82	0.7	0.88
	6	0.97	0.95	0.97	0.96	0.92		6	0.66	0.76	0.73	0.75	0.48
	7	0.95	0.96	0.96	0.96	0.95		7	0.4	0.75	0.8	0.77	0.84
	8	0.9	0.95	0.95	0.96	0.94		8	0.82	0.8	0.77	0.84	0.82
	9	0.99	0.94	0.97	0.84	0.81		9	0.89	0.79	0.92	0.75	0.77
	10	0.94	0.93	0.92	0.92	0.92		10	0.35	0.6	0.72	0.58	0.45
	11	0.87	0.93	0.92	0.93	0.91		11	0.68	0.83	0.86	0.83	0.84
	12	0.95	0.96	0.97	0.97	0.94		12	0.84	0.89	0.95	0.91	0.92
	13	0.97	0.98	0.98	0.97	0.96		13	0.97	0.96	0.92	0.9	0.93
	14	0.95	0.96	0.96	0.92	0.96		14	0.87	0.89	0.86	0.89	0.85

Table 31: Proportion of explained variability, R^2, when an estimated size explains twelve accounting items by industry and by year. Second Table. Numbers from 1 to 14 identify industries: 1, Building Materials; 2, Metallurgy; 3, Paper and Packing; 4, Chemicals; 5, Electrical; 6, Industrial Plants; 7, Machine Tools; 8, Electronics; 9, Motor Components; 10, Clothing; 11, Wool, 12, Miscellaneous Textiles, 13, Leather; 14, Food Manufacturers.

S	NW	W	I	D	C	CA	FA	TA	CL	TC	
0.366						1983, Building Materials					S
0.325	0.357										NW
0.350	0.331	0.360									W
0.388	0.359	0.398	0.472								I
0.350	0.304	0.345	0.383	0.358							D
0.358	0.296	0.340	0.380	0.358	0.387						C
0.344	0.314	0.342	0.386	0.338	0.335	0.339					CA
0.341	0.337	0.335	0.369	0.324	0.327	0.324	0.359				FA
0.344	0.323	0.339	0.379	0.334	0.333	0.334	0.338	0.336			TA
0.353	0.303	0.341	0.383	0.356	0.370	0.339	0.327	0.336	0.369		CL
0.345	0.350	0.348	0.382	0.327	0.317	0.336	0.356	0.343	0.321	0.369	TC
0.186						1983, Clothing					S
0.182	0.236										NW
0.176	0.172	0.205									W
0.214	0.221	0.199	0.277								I
0.188	0.194	0.168	0.238	0.236							D
0.189	0.181	0.178	0.226	0.206	0.220						C
0.184	0.199	0.170	0.225	0.202	0.195	0.200					CA
0.182	0.196	0.197	0.214	0.193	0.188	0.184	0.262				FA
0.180	0.193	0.174	0.216	0.195	0.189	0.191	0.202	0.190			TA
0.191	0.175	0.186	0.233	0.213	0.215	0.198	0.209	0.199	0.235		CL
0.181	0.227	0.171	0.220	0.194	0.182	0.199	0.197	0.193	0.176	0.226	TC
0.620											S
0.578	0.745					1983, Food					NW
0.588	0.682	0.724									W
0.656	0.675	0.660	0.777								I
0.575	0.585	0.578	0.636	0.570							D
0.587	0.589	0.595	0.651	0.571	0.608						C
0.576	0.620	0.591	0.660	0.569	0.584	0.597					CA
0.612	0.764	0.769	0.699	0.612	0.629	0.634	0.904				FA
0.579	0.647	0.626	0.662	0.570	0.587	0.596	0.692	0.615			TA
0.579	0.586	0.588	0.652	0.563	0.591	0.579	0.633	0.586	0.589		CL
0.592	0.736	0.697	0.688	0.594	0.602	0.628	0.785	0.660	0.601	0.745	TC

Figure 29: Typical Σ matrices. Items having only positive cases.

Ratio	1983		1984		1985		1986		1987	
	skew	kurt	skew	kurt	skew	kurt	skew	kurt	skew	kurt
D/C	5.7	44.5	5.6	48.3	14.0	255.0	15.0	261.0	4.8	35.5
C/D	18.7	381.1	20.5	461.0	22.5	535.0	14.3	249.0	12.4	198.0
CA/CL	2.0	5.8	4.2	29.2	5.0	38.4	11.6	205.0	4.8	43.6
CL/CA	19.6	415.0	4.2	34.5	3.0	22.9	9.2	135.0	21.6	506.0
C/I	9.9	139.0	9.8	132.0	24.2	592.0	15.8	268.0	21.3	491.0
I/C	6.5	71.5	3.3	18.0	8.9	119.0	9.2	119.0	7.9	91.3
Q/CL	0.0	14.2	3.1	21.8	4.0	31.6	7.3	99.3	5.5	64.7
CL/Q	11.9	192.0	2.5	97.2	-21.0	486.0	6.6	97.8	-18.7	418.0
W/N	18.9	399.0	1.5	4.1	1.5	4.0	1.4	3.3	1.6	4.4
N/W	11.7	199.0	2.2	12.6	7.6	112.0	7.8	118.0	8.9	116.0
S/TA	7.0	72.3	13.6	238.0	13.3	241.0	6.3	51.7	8.5	95.3
TA/S	8.8	111.0	19.5	436.0	7.6	89.0	15.9	336.0	3.4	22.3
S/FA	9.8	105.0	21.4	486.0	10.4	115.0	9.2	96.1	9.6	103.0
FA/S	2.3	9.4	15.9	317.0	8.9	125.0	5.1	50.8	5.9	60.9
S/NW	16.4	310.0	13.2	188.0	23.0	549.0	16.1	324.0	16.5	207.0
NW/S	1.6	5.2	4.4	38.6	2.5	17.6	12.0	223.0	4.7	39.8
S/I	11.7	168.0	12.0	189.0	17.0	306.0	23.5	571.0	21.3	481.0
I/S	8.9	132.0	17.1	358.0	17.9	388.0	9.8	179.0	1.4	4.5
EB/TA	2.0	8.3	2.4	11.2	1.4	3.6	1.9	7.3	1.2	2.6
TA/EB	16.6	291.0	12.2	172.0	11.1	173.0	7.5	76.0	18.3	374.0

Table 32: The skewness and kurtosis of ratios selected so as to avoid constraints. All groups. 1^{st} table. $Q = CA - I$; $TD = DEBT + CL$.

Ratio	1983		1984		1985		1986		1987	
	skew	kurt	skew	kurt	skew	kurt	skew	kurt	skew	kurt
EB/NW	19.9	425.0	17.5	317.0	16.6	322.0	8.4	110.0	16.4	330.0
NW/EB	15.7	256.0	17.6	369.0	12.7	227.0	7.0	71.9	23.4	567.0
D/I	10.4	127.0	16.6	329.0	21.6	578.0	24.4	618.0	19.9	423.0
I/D	10.7	128.0	11.8	166.3	16.0	298.0	10.0	126.8	7.5	74.6
W/I	15.8	268.0	9.3	116.2	7.9	79.0	13.6	204.5	14.4	219.3
I/W	8.6	89.7	4.6	28.1	4.5	31.2	3.1	12.7	3.3	14.0
EB/FA	9.4	117.0	11.2	169.8	10.4	154.7	9.0	107.0	8.5	98.8
FA/EB	12.1	154.0	9.8	126.0	11.6	155.7	7.7	83.8	16.8	320.8
S/N	12.7	189.0	12.3	174.0	7.8	75.6	8.7	89.3	11.1	159.8
N/S	1.7	5.8	1.1	1.3	3.7	31.0	2.6	16.9	3.3	27.9
EB/N	9.1	92.5	11.7	177.7	5.2	42.0	3.4	14.6	5.2	45.0
N/EB	15.9	268.0	9.0	130.7	6.0	46.1	8.0	84.2	13.4	212.4
NW/N	8.4	90.9	5.0	34.1	3.5	16.1	3.9	23.0	4.3	29.1
N/NW	6.4	56.4	4.8	42.7	18.2	359.8	4.2	24.5	8.2	112.6
W/TA	1.3	4.9	1.5	5.8	1.4	4.3	1.8	7.2	1.5	6.4
TA/W	9.5	107.0	11.5	152.9	5.7	50.6	4.2	27.0	3.7	19.7
DB/NW	14.5	228.0	10.9	140.0	8.6	90.7	7.5	78.0	5.4	37.4
NW/DB	13.3	199.8	17.6	332.7	12.0	184.9	11.0	145.6	16.1	295.7
DB/S	2.5	8.4	6.0	64.2	15.3	273.4	8.3	98.7	10.3	155.8
S/DB	10.0	120.4	16.3	293.6	12.1	178.2			10.2	113.2

Table 33: The skewness and kurtosis of ratios selected so as to avoid constraints. All groups. 2^{nd} table. $Q = CA - I$; $TD = DEBT + CL$.

Number	Industrial group	1983	1984	1985	1986	1987
Log items						
1	BUILDING MATERIALS	-0.31	-0.16	-0.29	-0.41	-0.34
2	METALLURGY	0.85	0.80	0.88	0.80	0.88
3	PAPER AND PACKING	0.65	-0.29	-0.40	-0.20	-0.12
4	CHEMICALS	-0.04	-0.36	-0.41	-0.51	-0.41
5	ELECTRICITY	0.22	-0.16	-0.45	-0.63	-0.65
6	INDUSTRIAL PLANTS	-0.76	-0.34	0.07	0.54	-0.12
7	MACHINE TOOLS	-0.88	-0.98	-1.21	-1.02	-1.20
8	ELECTRONICS	0.72	0.54	0.33	0.31	0.10
9	MOTOR COMPONENTS	0.84	0.67	0.48	0.15	0.44
10	CLOTHING	-1.81	-1.78	-2.03	-2.02	-1.97
11	WOOL	-1.63	-2.09	-1.36	-1.47	-1.16
12	MISC. TEXTILES	1.05	1.19	1.77	1.56	1.96
13	LEATHER	-1.12	1.25	1.05	1.18	0.67
14	FOOD MANUFACTURERS	1.53	1.25	1.18	1.36	1.44
Log residuals						
1	BUILDING MATERIALS	-0.72	-0.69	-0.85	-0.91	-1.08
2	METALLURGY	2.40	2.87	3.17	3.19	2.96
3	PAPER AND PACKING	-0.04	1.19	0.18	0.05	-0.36
4	CHEMICALS	-0.22	-0.17	0.08	0.24	0.31
5	ELECTRICITY	-0.54	0.27	0.19	0.48	-0.51
6	INDUSTRIAL PLANTS	-0.49	-0.59	-0.51	-0.65	0.70
7	MACHINE TOOLS	1.85	0.34	0.64	-0.58	-1.21
8	ELECTRONICS	0.10	-0.01	0.40	0.23	0.33
9	MOTOR COMPONENTS	-1.27	-1.24	-1.00	-0.65	-0.53
10	CLOTHING	0.09	-0.02	-0.48	-0.23	-0.44
11	WOOL	0.33	-0.50	-0.43	-0.66	-0.48
12	MISC. TEXTILES	0.12	-0.22	-1.02	-0.69	0.44
13	LEATHER	-1.71	-1.53	-0.84	-0.95	-1.08
14	FOOD MANUFACTURERS	-0.05	-0.06	0.11	0.57	0.62

Table 34: Scores ranking the spread of log items and log residuals for industrial groups.

Company		1983	1984	1985	1986	1987
		Mean-adjusted log size				
26	CAMPBELL FROZEN FOODS LTD	-0.24	-0.24	-0.25	-0.15	-0.16
51	LOVELL (G.F.) PLC	-1.09	-1.11	-1.18	-1.15	-1.16
55	MAUNDER (LLOYD) LTD	-0.26	-0.23	-0.27	-0.31	-0.27
60	NESTLE HOLDINGS (U.K.) PLC	0.975	1.005	0.998	0.991	0.988
62	OVERSEAS FARMERS'CO-OP FE	-0.88	-1.01	-1.05	-0.94	-0.87
86	UNITED BISCUITS (HOLDINGS)	1.397	1.447	1.439	1.447	1.448
		Funds Flow to Total Debt				
26	CAMPBELL FROZEN FOODS LTD	0.478	0.524	0.866	0.599	0.419
51	LOVELL (G.F.) PLC	0.043	0.122	0.159	0.387	-0.01
55	MAUNDER (LLOYD) LTD	0.179	0.207	0.161	0.157	0.206
60	NESTLE HOLDINGS (U.K.) PLC	0.223	0.301	0.338	0.442	0.477
62	OVERSEAS FARMERS'CO-OP FE	0.204	0.213	0.158	0.057	0.033
86	UNITED BISCUITS (HOLDINGS)	0.323	0.274	0.329	0.332	0.420
		EBIT to Net Worth				
26	CAMPBELL FROZEN FOODS LTD	0.218	0.232	0.227	0.100	0.064
51	LOVELL (G.F.) PLC	-0.01	-0.03	0.020	0.138	-0.12
55	MAUNDER (LLOYD) LTD	0.026	0.073	0.053	0.091	0.227
60	NESTLE HOLDINGS (U.K.) PLC	0.152	0.310	0.366	0.538	1.336
62	OVERSEAS FARMERS'CO-OP FE	0.192	0.299	0.203	0.097	0.057
86	UNITED BISCUITS (HOLDINGS)	0.394	0.372	0.307	0.325	0.338
		Sales to Net Worth				
26	CAMPBELL FROZEN FOODS LTD	3.37	3.01	2.67	1.51	1.51
51	LOVELL (G.F.) PLC	3.26	4.67	3.19	3.30	3.81
55	MAUNDER (LLOYD) LTD	10.27	12.70	13.81	17.54	16.22
60	NESTLE HOLDINGS (U.K.) PLC	5.60	4.56	5.21	6.02	9.55
62	OVERSEAS FARMERS'CO-OP FE	9.22	13.70	11.04	15.96	31.90
86	UNITED BISCUITS (HOLDINGS)	5.26	5.51	4.36	4.02	3.74

Table 35: Some ratios and an estimated size, during the period 1983-1987 for nine firms.

Bibliography

[1] J. Aitchison and J. Brown. *The Lognormal Distribution.* Cambridge University Press, 1957.

[2] P. Barnes. Methodological implications of non-normally distributed financial ratios. *Journal of Business, Finance and Accounting*, pages 51–62, Spring 1982.

[3] W. Beaver. Financial ratios as predictors of failure. *Journal of Accounting Research, Suplement. Empirical Research in Accounting: Select Studies*, pages 71–111, 1966.

[4] R. Berry and S. Nix. Regression analysis vs. ratios in the cross-sectional analysis of financial statements. *Accounting and Business Research*, 21(82):107–117, 1991.

[5] R. Bird and A. McHugh. Financial ratios - an empirical study. *Journal of Business Finance and Accounting*, pages 29–45, Spring 1977.

[6] P. D. Bougen and J. Drury. U.k. statistical distributions of financial ratios, 1975. *Journal of Business, Finance and Accounting*, 1(7):39–47, 1980.

[7] W. Buijink and M. Jegers. Cross-sectional distributional properties of financial ratios in belgian manufacturing industries: Some empirical evidence. Technical report, University of Antwerp, Belgium, 1984.

[8] K. Chen and A. Shimerda. An empirical analysis of useful financial ratios. *Financial Management*, pages 11–26, Spring 1981.

[9] R. Cook. Detection of influential observations in regressions. *Technometrics*, 19:15–18, 1977.

[10] E. Deakin. Distributions of financial accounting ratios: Some empirical evidence. *Accounting Review*, pages 90–96, January 1977.

[11] R. Eisenbeis. Pitfalls in the application of discriminant analysis in business, finance and economics. *Journal of Finance*, 32(3):875–899, June 1977.

[12] E. Elton and M. Gruber. Homogeneous groups and the test of economic hypothesis. *Journal of Financial and Quantitative Analysis*, 4:581–602, 1970.

[13] M. Ezzamel and C. Mar-Molinero. The distributional proprieties of financial ratios in u.k. manufacturing companies. *Journal of Business, Finance and Accounting*, 17(1):1–29, 1990.

[14] M. Ezzamel, C. Mar-Molinero, and A. Beecher. On the distributional properties of financial ratios. *Journal of Business, Finance and Accounting*, 14(4):463–481, 1987.

[15] G. Foster. *Financial Statement Analysis.* Prentice-Hall, 1986.

[16] T. Frecka and W. Hopwood. The effect of outliers on the cross-sectional distributional properties of financial ratios. *The Accounting Review*, pages 115–128, January 1983.

[17] R. Gibrat. *Les Inegalites Economiques.* Librairie du Recueil Sirey, 1931.

[18] J. Horrigan. Some empirical bases of financial ratio analysis. *Accounting Review*, pages 558–568, July 1965.

[19] J. Horrigan. Methodological implications of non-normally distributed financial ratios: A comment. *Journal of Business Finance and Accounting*, pages 683–689, Winter 1983.

[20] Y. Ijiri and H. Simon. *Skewed Distributions and the Size of Business Firms.* North-Holland, Amsterdam, 1977.

[21] R. Jensen. A cluster analysis study of financial performance of selected business firms. *The Accounting Review*, pages 36–55, January 1971.

[22] B. Lev and S. Sunder. Methodological issues in the use of financial ratios. *Journal of Accounting and Economics*, pages 187–210, December 1979.

[23] S. Mcleay. The ratio of means, the mean of ratios and other benchmarks. *Finance, Journal of the French Finance Society*, 7(1):75–93, 1986.

[24] S. Mcleay. Student's t and the distribution of financial ratios. *Journal of Business, Finance and Accounting*, 13(2):209–222, 1986.

[25] M. O'Connor. On the usefulness of financial ratios to the investor in common stock. *The Accounting review*, pages 339–352, April 1973.

[26] J. P. Royston. An extension of the shapiro and wilk's test for normality to large samples. *Applied Statistics*, 31(2):115–124, 1982.

[27] G. Snedecor and W. Cochran. *Statistical Methods.* Iowa State University Press, 9^{th} edition, 1965.

[28] J. Steindl. *Random Process and the Growth of Firms: A Study of the Pareto Law.* Hafner, New York, 1965.

[29] P. Sudarsanam and R. Taffler. Industrial classification in u.k. capital markets: A test of economic homogeneity. *Applied Economics*, 17:291–308, 1985.

[30] G. Whittington. Some basic properties of accounting ratios. *Journal of Business, Finance and Accounting*, 7(2):219–232, 1980.