The Foundations of the Financial Ratio Measurement

Duarte Trigueiros

Faculty of Business Administration,
University of Macau,
Po Box 3001, Macau

Fax: 00 853 831 694
e-mail: fbadt@umac.mo

The British Accounting Association National Conference, Birmingham, March 1997.

The Foundations of the Financial Ratio Measurement

Abstract

The paper proposes a theoretical foundation for the financial ratio measurement, showing that the multiplicative character of ratio components is a necessary condition for valid ratio usage, not just an assumption supported by evidence. Also, by assuming that firm size is a measurable statistical effect, the paper offers a better defined discussion of limitations of ratios. Thus a well-known limitation, non-proportionality, is re-assessed and a new, potentially serious limitation is described.

The paper has two parts, one where ratio components are viewed as deterministic variables and the other were they are random. Such approach allows an easier understanding of the expected ratio before the generalisation to encompass randomness takes place. In the second part, ratio components are compared to 'random effects' models. It is shown that, in this case, a specific type of exponential Brownian motion leads to ratios which will not necessarily drift.

The Foundations of the Financial Ratio Measurement

Introduction

Accounting academics have often considered the widespread use of ratios in financial analysis as a somehow intriguing practice. As early as 1965, Horrigan noticed that financial ratios were referred to in text books in almost apologetic tones as though their expected utility were low. He tried to dissipate such doubts by describing statistical characteristics of some widely used ratios, concluding that they may be useful after all.

In response to Horrigan's optimistic view, the late sixties uncovered promising applications of ratios. Beaver (1967) and Altman (1968), for example, showed that ratios have the potential to help predicting bankruptcy. A few years later, however, the negative tone returned as authors such as Deakin (1976) noticed that statistical distributions found in ratios may vary widely. This prompted Frecka & Hopwood (1983) and other authors to propose *ad hoc* techniques (such as applying transformations and then trimming or winsorising outliers) to deal with ratios. Those and similar suggestions reflect the widespread belief that no general rules are applicable to the ratio method.

Adding to this belief, Lev & Sunder (1979) questioned whether the use of ratios is motivated by well founded considerations or just by tradition. These authors claimed that almost all of the assumptions required for valid ratio analysis are likely to be violated in practice. A more balanced critique followed (Whittington, 1980) uncovering cases where ratios seem not to be up to the task but distinguishing between normative and predictive applications, considering the former as acceptable.

Both Lev & Sunder (1979) and Whittington (1980) stressed the fact that valid measurement using ratios requires proportionality between components. Since such assumption seems to be too restrictive, these authors advocated a regression rather than a ratio approach. Barnes (1982) went one step further, suggesting that difficulties caused by skewed distributions in ratios also stem from non-proportionality. He showed that the use of regressions or similar functional forms instead of ratios should, in this case, eliminate both the problem of non-proportionality and that of skewness in distributions.

A striking feature of the above-mentioned research is the small impact it had, both in financial analysts' practice and in the way empirical research is carried out. One reason may be some lack of awareness of related developments in sciences such as Industrial Economics. Authors seem to pre-suppose that accounting data is an especial case, too complex to allow simple, unifying explanations. As a consequence, interpretation of empirical results is avoided or, when attempted, assumptions are timid, leading to conclusions which are too general to be useful.

For example, in spite of insisting that ratios are aimed at removing size, to date the literature on ratio analysis did not come out with a definition of what size may be. However, if ratios are indeed aimed at removing size, no progress can be made unless specific definitions of firm size are tested against the ratio method.

A major step towards a more focused research was taken by McLeay (1986*a*; 1986*b*) who suggested that, rather than drawing inferences from trimmed means, data should be left unadjusted and appropriate models should be developed to fit characteristics of accounting data. This author then presented a few such models, amongst them a multiplicative[1] model. More recently, Tippitt (1990) proposed an inductive methodology to study the distribution of ratios where components are also assumed to be multiplicative and Trigueiros (1995) and Tippet & Whittington (1995) provided empirical evidence on the multiplicative character of ratio components. The former author also offered a simple explanation for anomalies observed in the distribution of ratios.

Given the above, it seems as though the time is ripe to develop a theoretical foundation for financial ratio measurement, in the light of which ratios may be viewed as simple tools governed by simple rules. This paper, while proposing one such foundation, reinforces intuitions of authors such as McLeay (1986*a*), Tippett (1990) and Trigueiros (1995), showing that the multiplicative character of components is a condition for valid ratio usage, not just a reasonable assumption supported by empirical observation.

Also, by assuming that firm size is a measurable statistical effect the paper is able to offer a more distinct discussion of limitations of ratios. Thus a well-known limitation, non-proportionality, is re-assessed and a new, potentially serious limitation is described. The study also suggests methodologies which may lead to more accurate ratios.

The paper is organised in two parts, one where ratio components are viewed as deterministic variables and the other were they are random. Such approach allows overall characteristics of ratios to be easily understood before being generalised to encompass randomness. Also, since determinism is the limiting case where volatility is negligible, the lower bound on complexity thus obtained helps interpreting the theoretical developments presented.

Validity of Deterministic Ratios

Financial analysis is just one of the many task domains where ratios are routinely used. There is a multitude of other cases where the usefulness of ratios is evident. Scales in maps and reduced models, for example, are ratios measuring the number of times a model is smaller than in reality. Speed, flow, or pressure are also ratios, measuring the average change in the numerator per unit change in the denominator. Derivatives, *i.e.*, rates of change, are also ratios of two small changes.

Most of the ratios used in financial analysis are similar to scales: sales margin, for instance, measures the number of times profits is smaller than sales; interest cover is the number of times earnings is bigger than interest; the liquidity ratio is the number of times current assets exceeds current liabilities, and so on.[2]

Consider two variables Y and X and suppose that there is a function $Y = f(X)$ describing Y in terms of X (figure 1). A change in X, say, from A to B, is $X_B - X_A$ and the corresponding change in Y is $Y_B - Y_A$. The ratio, r, is

$$r = \frac{Y_B - Y_A}{X_B - X_A}.$$

From the definition of derivative, or by expanding $f(X)$ in Taylor series around a point between A and B, it is clear that r is the linear approximation to $f'(X)$, the derivative of $f(X)$. Ratios, in fact, are the simplest formulation aimed at measuring how much a variable changes when changes in a related variable take place.

When, between A and B, $f(X)$ is a straight line, then the ratio exactly equals $f'(X)$. In this case, the rate of change of Y with X is a constant value, not a variable, and it measures the slope of the chord AB. When, between A and B, $f(X)$ is not a straight line, r equals the average $f'(X)$. The more convex $f(X)$ is, the worst r approximates $f'(X)$. r is called *first order difference quotient*. In practice derivatives are calculated using first order difference quotients measured over small intervals.

Logarithms perform, using differences, a task similar to that of ratios, *i.e.*, measurement on a logarithmic scale is a rate of change. In logarithms, changes, say, from one million to two millions or from one billion to two billions will have the same value whose meaning will be *twice*. When considering ratio validity requisites, this affinity between ratios and logarithms is stretched even farther.

Lev & Sunder (1979), Whittington (1980) and others have remarked that ratios are valid where the relationship between the numerator and the denominator is a straight line passing through the origin of co-ordinates, *i.e.*,

$$\frac{Y}{X} = \text{Constant}.$$

By stating that the expected ratio should remain constant no matter changes in components, such 'traditional' definition aims at ensuring the removal of size from measurement. However, it may also be revealing to observe, not only the function $f(X)$ required to remove size, but also $f'(X)$, *i.e.*, how changes in Y relate to those in X. The above-mentioned equality between constant ratios and derivatives may be written

$$\frac{Y}{X} = \frac{dY}{dX}$$

where dY, dX are related changes observed in Y and X. As a consequence,

$$\frac{dY}{Y} = \frac{dX}{X}. \tag{1}$$

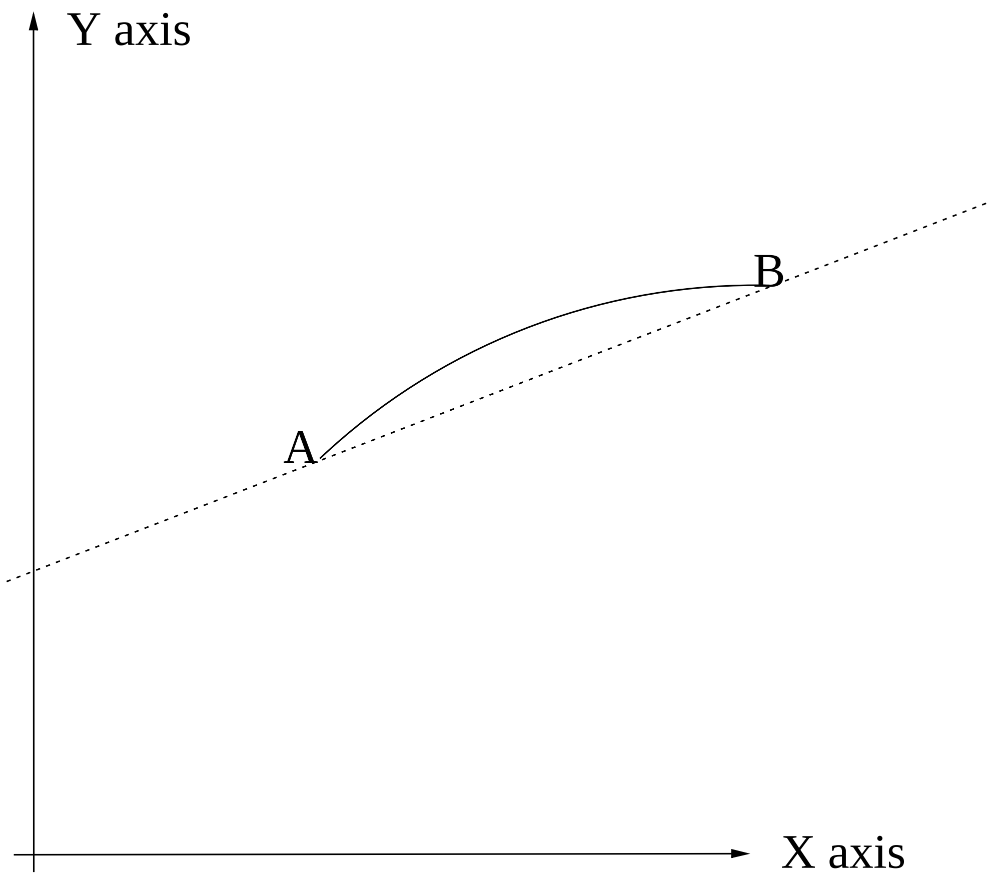

Figure 0

When the slope goes through the origin, a ratio is a derivative

Figure 1: Derivatives measure the steepness of a slope. Ratios, as linear approximations to derivatives, do the same when slopes are straight lines.

Equality (1) suggests that conditions for ratio validity are more explicit than those conveyed by the 'traditional' definition. In the following, two such conditions, scale-invariance and exponential components, are separately analysed.

In the first place, (1) shows that, in order for ratios to be valid, expected percentage changes leading to numbers in the same report must be similar. For instance, when comparing reports of two firms, a percentage difference of 200% in Current Assets (Current Assets in one firm is expected to be three times larger than in the other) should also be expected when comparing, say, Current Liabilities for the same two firms or, indeed, when comparing any other variable potentially useful as component of a ratio. In a time-series context (1) implies that any variable eligible as component of a ratio is expected to grow at the same rate. If, say, Sales grows 12% in a given year, then Earnings and other variables should also grow 12% during that year.

It should not be surprising that validity of ratios is conditional on the equality of percentage changes in variables. Since ratios are scales, they are valid only where scaling of data makes sense and this implies *scale-invariance* as a property of such data. Figure 2 provides an intuitive description of scale-invariance: triangle $A'B'C'$ is twice as large as triangle ABC whereas their angles are similar. Scaling, in such case, makes sense and, indeed, a simple ratio is enough to express the change whereby triangle ABC becomes similar in size to $A'B'C'$. However, scaling would not make sense when describing the change of triangle ABC into DEF because their angles are different.

Scale-invariant changes such as those leading to triangle $A'B'C'$ require that vectors $\{X_A, Y_A\}$, $\{X_B, Y_B\}$ and $\{X_C, Y_C\}$, when evolving along the dotted lines in figure 2, all obey the same function. Furthermore, such function must generate, in each vector, changes which are similar in percentage, as in (1). The change from A to D, for example, is five times the change from B to E or C to F. Should the three changes be the same in percentage, then triangle DEF would have been a scale-invariant replica of ABC.

The condition of an expected percentage rate for different numbers requires the existence of a common characteristic underlying such numbers. Size is the only characteristic which all numbers in a report are expected to share. In fact, it is commonplace that reports of large firms exhibit numbers which are many orders of magnitude larger than those in small firms' reports.[3] Percentage changes in variables should thus be viewed as reflecting differences in the size of two firms or, in a time-series context, as the growth rate of the firm.

Figure 2 also provides an intuitive representation of how size determines the expected magnitude of numbers in a report. Suppose that the X-axis measures size and the Y-axis is allowed to represent different variables. Specifically, let Sales be expected to be four times larger than size (corresponding to slope BB') and let Earnings be just 20% of size (slope AA'). Thus each variable has its specific *scale to size* which, in this example is $Y_B/X_B = Y_{B'}/X_{B'} = \cdots = 4$ for Sales and $Y_A/X_A = Y_{A'}/X_{A'} = \cdots = 0.2$ for Earnings.

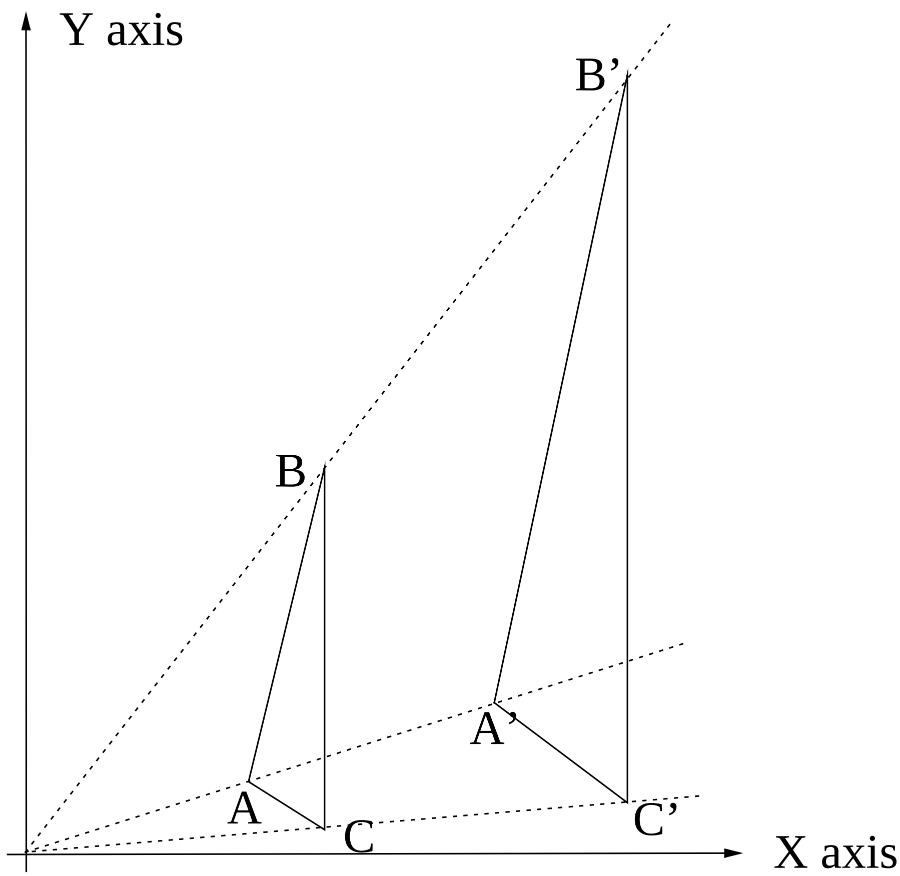

Figure 1
An example of scale invariance (ABC:A’B’C’)

Figure 2: An example of scale-invariance (triangles ABC and $A'B'C'$) together with a case where scale-invariance does not apply (triangle DEF).

The expected value of the ratio Earnings to Sales, 5%, stems from the division of the two scales to size involved (20% and 4 times). It is because the same change in size (differently scaled) is present in each component that its removal may take place.

The removal of size from the expected ratio is not enough to ensure validity. Size must also be removed from any *deviation* observed in relation to such expectation. If, say, an increment in 1% in ROE meant, for a small firm, just a slight improvement in relation to the industry whereas, for a large firm, the same increment meant a great achievement, then the measurement would be justly regarded as meaningless.

Such condition, not contemplated by the 'traditional' approach, is fulfilled where absolute changes in components of ratios are proportional to their actual values since, in such case, a ratio which is expected to be size-independent will remain size-independent even when a deviation from expectation occurs. If τ is the variable which drives changes dY or dX observed in component Y or X, then

$$\frac{dY}{d\tau} = s_y\, Y \quad \text{and} \quad \frac{dX}{d\tau} = s_x\, X$$

where s_y, s_x must be strictly independent of Y, X respectively. Or,

$$\frac{dY}{Y} = s_y\, d\tau \quad \text{and} \quad \frac{dX}{X} = s_x\, d\tau.$$

This condition is known as the 'Law of Proportionate Effect' or simply as the 'Gibrat Law'. It is in the origin of the 'multiplicative' family of variables, of which the exponential function of τ *e.g.*,

$$Y = {\scriptstyle Y_0}\, e^{s_y\, \tau} \quad \text{and} \quad X = {\scriptstyle X_0}\, e^{s_x\, \tau}. \tag{2}$$

is the simplest instance. ${\scriptstyle Y_0}$, ${\scriptstyle X_0}$ are arbitrary constant magnitudes. When τ is the same for both components of ratios, then scale-invariance is verified where $s_y = s_x$.

In summary, validity of ratios requires percentage changes in components to be, not only similar for different components, but also independent of their actual values. Whereas the first condition, scale-invariance, demands a specific type of relationship between components, this second condition is a constraint which is imposed on the behaviour of individual components.

Notice that (2) refers to continuous rates of change s_y and s_x whereas (1) equates two effective[4] rates. In fact, Y and X evolve similarly to amounts ${\scriptstyle Y_0}$, ${\scriptstyle X_0}$ earning, during τ, continuous compound interest s_y, s_x.

When Y and X are random or when dY and dX are not infinitesimal,[5] the distinction between continuous and effective rates must be taken into account. A robust formulation of scale-invariance should equate continuous rather than effective rates, *i.e.*,

$$\mathcal{E}\left[\log \frac{y + dy}{y}\right] = \mathcal{E}\left[\log \frac{x + dx}{x}\right]$$

(where $\mathcal{E}$ denotes expectation) or, abridged,

$$d(\log Y) = d(\log X). \tag{3}$$

In fact, (3) encompasses (1) but retains its meaning under broader circumstances. Equality (3) is also interesting in that, contrarily to (1), it may lead to constant expected ratios when components are stochastic variables.

In a time-series context, the interpretation of (2) is straightforward. τ is a clock measuring a time-sequence starting when $Y = \mathrm{Y}_0$, $X = \mathrm{X}_0$. These initial values should be viewed as measuring the actual size of the firm, differently scaled according to 'scales to size' specific to each variable. s_y, s_x are real growth rates observed in Y or in X. These growth rates may vary widely with time but they are expected to be similar for different variables since they mostly reflect percentage changes in the size of the firm. Deviations observed in s_y or s_x in relation to the growth rate of the firm, s, may be expressed as a difference $n_y = s - s_y$ or $n_x = s - s_x$.

Supposing that, in a given year, Sales is £1,000 and Earnings is £100, then profitability is 10%. If Sales grows s (the same as the firm) but Earnings grows only $s - n$, then, in the following year,

$$\text{Earnings} = 100\, e^{(s-n)\tau} = 100\, e^{s\tau}\, e^{-n\tau} \quad \text{and} \quad \text{Sales} = 1,000\, e^{s\tau}$$

and the ratio decreases by $e^{-n\tau}$ irrespective of s. However small or large growth may be during a given year, ratios remain comparable for different years.

In a cross-section context τ should ideally measure standardised size relative to industry expectation. In practice, τ measures the number of standard deviations separating each realisation of Y or X from the corresponding industry expectation Y_0, X_0. As for s_y, s_x, they should also be similar for different variables, denoting the dispersion of sizes. Industries where firms range from the very small to the very large exhibit large s whereas those where size is more regular exhibit small s.

For example, in a given industry the median Sales is £1,000 and Earnings is £100 whereas, for report j, both Sales and Earnings are τ_j standard deviations above or below these 'normal' values, *i.e.*,

$$\text{Earnings}_j = 100\, e^{s\tau_j} \quad \text{and} \quad \text{Sales}_j = 1,000\, e^{s\tau_j}.$$

Now consider k, another report reflecting a more profitable situation, where the same volume of Sales generates Earnings n standard deviations above τ_j, *i.e.*,

$$\text{Earnings}_k = 100\, e^{s(\tau_j+n)} = \text{Earnings}_j\, e^{sn} \quad \text{and} \quad \text{Sales}_k = \text{Sales}_j.$$

The percentage difference in profitability between j and k, e^{+sn}, is independent of τ_j. Thus, in (2) or in similar functions, however small or large firms are, differences in ratios are comparable across firms.

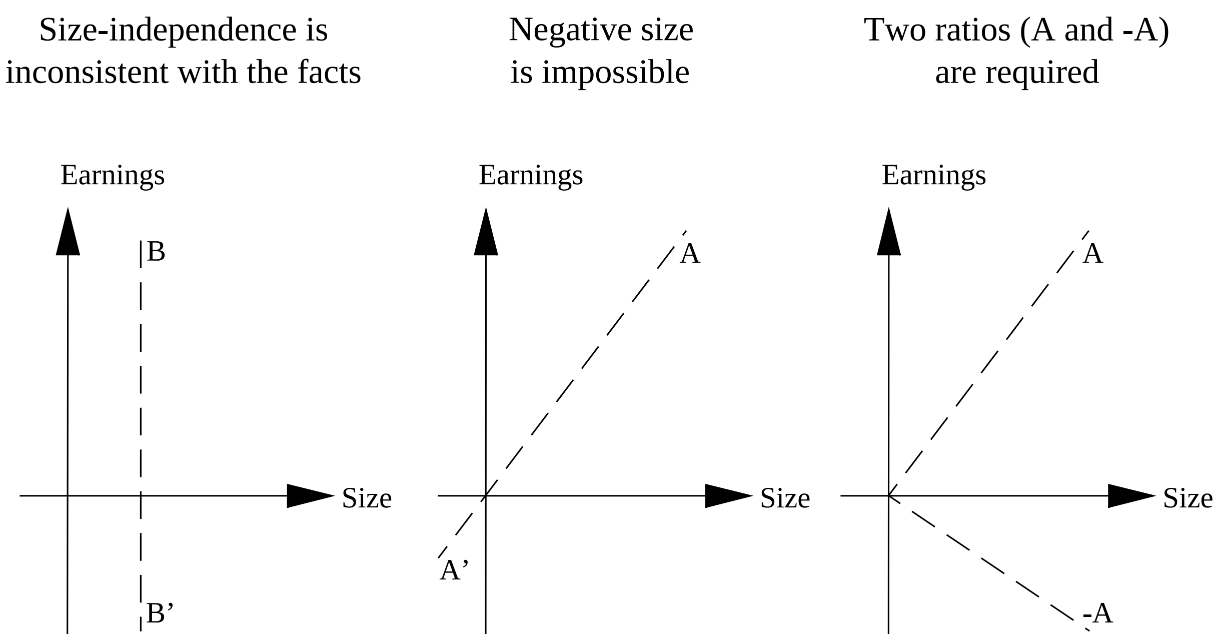

Figure 3: The modelling of the relationship between Earnings and size requires two ratios (slopes A and $-A$) for positive and negative numbers respectively.

Moreover, the increment experienced by the ratio is also independent of Y_0, X_0. Independence from industry norms, initial values, targets, benchmarks and so on, should not be regarded as a minor condition for the validity of ratios as practitioners make decisions based on measurement of deviation from these.

Non-exponential components lead to size- or norms-dependent ratios. Variables proportional to size, (*e.g.*, Earnings $= 4\, s\tau$ and Sales $= 40\, s\tau$), satisfy the two ‘traditional’ validity requisites, linearity and convergence. Moreover, apparently, they also obey (1), at least if τ is not allowed to approach zero. However, deviations in ratios are influenced by norms and by size. Where a gain $d\tau = n$ is observed in Earnings, the ratio would increase by $10\%\, n/\, s\tau$. A value of $n = 1$ means 5% extra profitability for the second firm in a rank whereas, for the 10th firm it would mean just 1%. It may also be worth noticing that any variable that becomes zero when τ is zero, is inadequate to describe growth since an ‘amount’ of zero money units cannot grow.

A final issue worth addressing is the way negative components of ratios should be manipulated. In the case of random components, for instance, it may be asked whether ratios should be viewed as one distribution, or rather as a juxtaposition of two distributions, one for positive and the other for negative realisations. Practitioners, when assessing industry norms, put aside negative numbers. However, in other cases *e.g.*, in the building of Z-scores, the use of a unique ratio seems more appealing.

Figure 3 shows that, when using a unique ratio, the assumption implicitly accepted is that, either firm size is allowed to be negative (which is impossible) or else, if size must be positive, then variables such as Earnings or Working Capital where negative numbers occur must be independent of size. In fact, the only line in figure 3 where size is always positive is BB' denoting size-independence. Since realisations of accounting variables, either positive or negative, are indeed size-dependent (otherwise ratios would not be necessary anyway), line BB' contradicts facts.

As a consequence, although the consideration of a unique ratio or distribution may bear some appeal, when choosing to use one ratio solely practitioners should be aware that such ratio will not remove size.

Limitations of Deterministic Ratios

Limitations of ratios may better be discussed by studying how characteristics of accounting data may restrict (1) and (2) in practice. Limitations stemming from non-exponential behaviour of variables (which lead to size- and norm-dependence), are better understood in a more specific context and they will be studied elsewhere.

Probably the two most obvious characteristics of accounting data which may limit the feasibility of (1) are (a) the presence of additive (*i.e.*, independent of size) terms and (b) differences in liquidity or other mechanisms able to prevent accounting variables from growing at similar percentage rates. Other, less pervasive causes may be (c) the existence of 'economies of scale' leading to non-linear changes (Whittington, 1980) and (d) the use of components belonging to different reports.

The presence, in components of ratios, of additive ('intercept') terms is a well-known limitation, generally presented as the main challenge to their valid use. Sales, for example, grows at a different rate from that of Earnings most likely because items such as Fixed Costs, Depreciation or Interest, all terms of Earnings, may not necessarily evolve as strict proportions of size.

Distortion introduced by additive terms is likely to be negligible in most instances. Due to their exponential character, accounting variables may easily attain values many times larger than their additive terms (which are size-independent). For example, where the denominator of a ratio is affected by an additive term δ_x then it evolve over time as

$$X = x_0 e^{s\tau} + \delta_x$$

instead of as in (2). Clearly, X tends to become much larger than δ_x. Moreover, as figure 4 depicts, this distortion is local, *i.e.*, it is significant only where variables are not much larger than their additive terms, asymptotically decaying for large sizes.

In a cross-section context, since observations belong to different firms, additivity requires industry-wide 'fixed costs'. Such overall costs must allow for the survival of small firms.[6] Therefore, where ratios are used for normative purposes, corrections of distortions caused by non-proportionality are likely to be required only when assessing deviations of small firms from

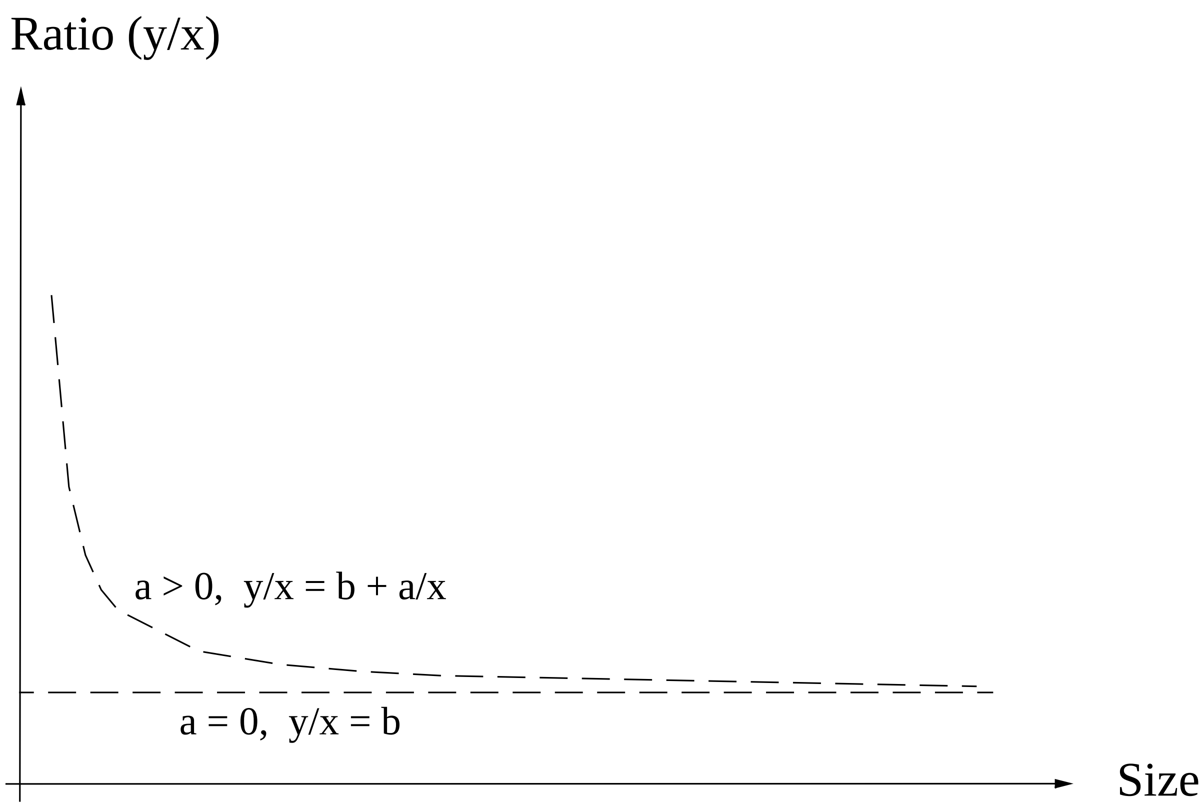

Figure 4
The local effect of an intercept term.

Figure 4: Ratio vs. size for different additive terms (A to D and no additive term (E).

industry norms. An industry-wide, significant distortion would be possible only in industries where firms have similar sizes.

In a time-series context, since observations belong to the same firm, additive terms such as Fixed Costs may be large comparative to ratio components and significant distortion is plausible.

Additive terms may be corrected, when measurable and stable, by introducing them back in (1). For example, Fixed Costs may be accounted for by writing

$$\frac{\text{Change in Operating Profits}}{\text{Operating Profits} + \text{Fixed Costs}} = \frac{\text{Change in Sales}}{\text{Sales}}$$

which is similar to using the ratio

$$\frac{\text{Operating Profits} + \text{Fixed Costs}}{\text{Sales}}$$

for measuring Sales Margin.

Differences in liquidity (or any other causes having the potential to delay or accelerate the growth of specific variables) create size-dependence in ratios. Requirements in Working Capital, for example, often grow faster than available Liquid Funds. When the denominator of a ratio grows faster than the numerator, large numbers form ratios which are smaller than expected. Conversely, when the numerator grows faster than the denominator, large numbers have their ratios distorted upwards.

As figure 5 depicts, this limitation may lead to serious distortion in measurement. Since the drift (per time- or size-unit) is proportional to the difference between rates of change in components, even a small difference introduces in ratios an exponential correlation with size, not just a disturbance with local effects only as in the case of additive terms.

This limitation is difficult to correct. Contrarily to additive terms, distortion is correlated with measurement, thus making it impossible to separate one from the other using the ratio solely. Although drifts in ratios are easily accounted for, there is the danger of mistakenly 'account for' features of the firm, such as a sustained extra efficiency.

Small drifts may be approximately accounted for by introducing in one of the terms of (1) a factor slightly larger or smaller than the unit whose effect is to diminish the correlation of the ratio with size. The current ratio, for instance, may become stable, even during fast growth, if one of its components is corrected, *e.g.*,

$$\frac{\text{Change in Current Assets}}{\text{Current Assets}} = b\,\frac{\text{Change in Current Liabilities}}{\text{Current Liabilities}}$$

which leads to the ratio

$$\frac{\text{Current Assets}}{(\text{Current Liabilities})^b}.$$

This heuristic method has the advantage of allowing easy parameter estimation and also of correcting 'economies of scale' affecting large sizes. However, it is unsatisfactory in that the

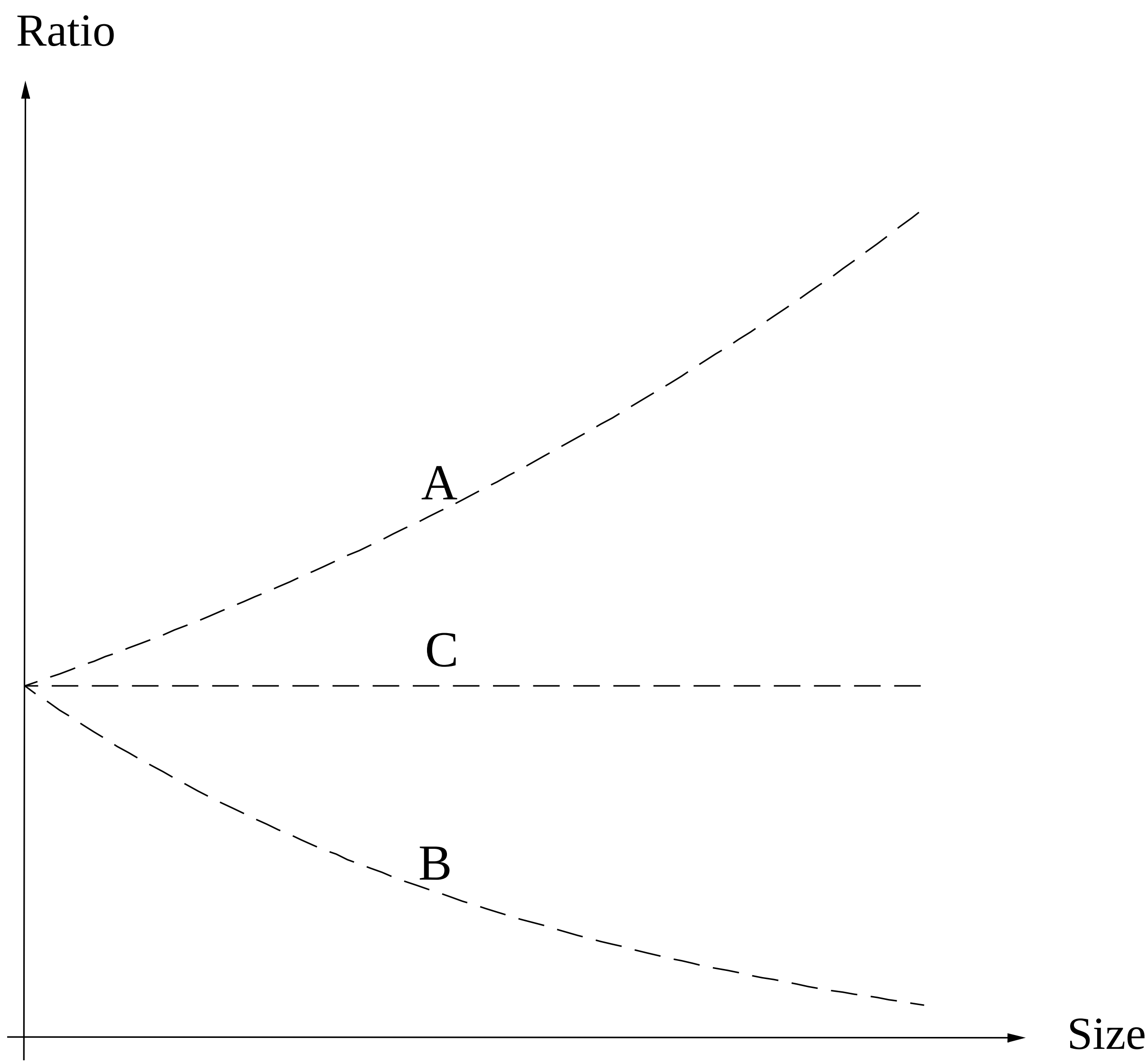

Figure 5

Size-dependence in ratios where the proportionate growth in the denominator is (A) smaller, (B) larger, or (C) the same as in the numerator.

Figure 5: Ratio vs. size when the percentage growth of the denominator is (A) smaller, (B) larger or (C) similar to that of the numerator.

expected ratio becomes different from that without correction, thus precluding comparison with other firms or industry norms.

Finally, the removal of size from measurement takes place only when the same size is present in both components. Valid candidates to components of ratios should thus belong to the same report. Ratios such as ROE or ROI, for which it is frequent to take the denominator from the previous year's report, will be correlated with size. Specifically, each case will exhibit a drift (relative to the ratio's expectation) proportional to the firm's actual rate of growth.

During periods of stable growth, ratios using lagged denominators may be fitted for time-series analysis as, in this case, the expected ratio is constant. The presence of size in the measurement may even be desired, *e.g.*, when monitoring growth. However, other cases exist where such presence may be spurious. EPS or PE ratios, for example, may be correlated with size because the number of shares in issue does not necessarily reflect the same size as that of Earnings.

Figure 6 summarises the above discussion. The intuitive representation introduced in figure 2 is used to illustrate limitations outlined above. Plot I shows scale-invariance. Sales is four times size (slope $BB'B''$) and Earnings is 20% of size (slope $AA'A''$). A small additive term subtracted from Earnings leads to plot II. Where Sales grows faster than Earnings the ratio decreases with size (plot III). Finally, ratios with lagged denominators are stable provided that growth is also stable (plot IV) but the expected ratio suffers o constant displacement in proportion to such growth.

It is clear from figure 6 that linearity and convergence (the 'traditional' definition) are necessary but not sufficient conditions for the validity of ratios. Where chords AB, $A'B'$ or $A''B''$ are also parallel with each other, then the ratio is constant but it may not necessarily remove size; where chords are also perpendicular to the X-axis, then components refer to the same size and the measurement is size-independent.

It is also clear that the removal of size requisite is different from that of constant expected measurement. Ratios may be constant but size-dependent (*e.g.*, for lagged denominators plus stable growth) or they may remove size but failing to remain constant (as for some stochastic processes which may be considered to describe components).

Validity of Random Ratios

When studying random characteristics of ratios, most authors have assumed that difficulties posed by their atypical behaviour are caused by distortions of normality. This is natural since the Normal distribution has dominated statistical practice. Most random events are nearly Normal or may be described as distortions of the Normal distribution. Indeed, any variable whose realisation stems from the additive contribution of many small influences tends to be normally distributed.

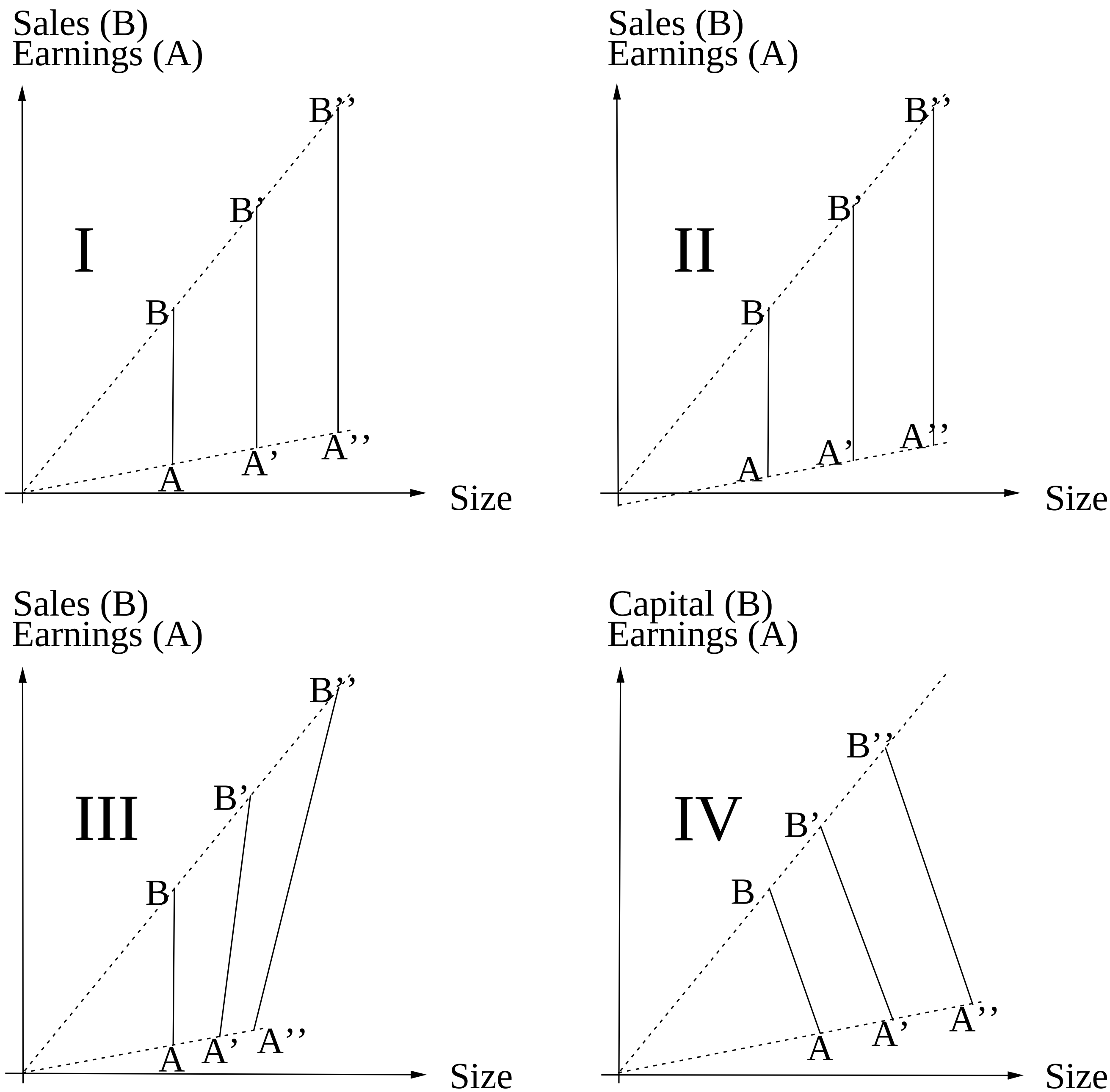

Figure 6

Scale-invariance (I); additive term (II); differences in proportionate growth (III); lagged denominators (IV).

Figure 6: Scale-invariance (I), additive terms (II), differences in percentage growth (III) and lagged denominators (IV).

However, accounting variables, as well as other events in Economics and Finance (such as individual income and wealth, stock prices, size or assets of firms) cannot be described as resulting from additive contributions. Any such realisation is generated under a multiplicative rather than an additive law of probabilities. For instance, each transaction leading to the total amount of Sales or each new investment in Fixed Assets is itself a random event and it contributes to the realisation, not by acting as an influence able to increase or decrease its likelihood, but by accumulation.[7]

Multiplicative variables tend to be lognormally rather than normally distributed (*i.e.*, their logarithms are nearly Normal[8]). They cannot be treated as distortions of normality as no distorting mechanism would be able to create, in additive events, the wide range of values generally found in multiplicative variables. It is frequent, for instance, that large cases in a lognormal sample be five hundred times larger than small cases. Such proportion has no counterpart in proportions observed, for example, between tall and short adults, a typically additive variable.

When the multiplicative character of accounting data is ignored, features which would otherwise be considered as commonplace (such as extreme positive skewness and outliers) risk to be taken as oddities. This is literally what has happened with some previous research. For example, most of the outliers often mentioned in relation to ratios probably are just a consequence of multiplicative skewness.

Most of the realisations of lognormal distributions concentrate in a small region while only a few extreme values spread out over a wide range (figure 7 on the right hand side). In the light of this, trimming, as advocated by authors such as Frecka & Hopwood (1983), is a cavalier practice. Moreover, since the mechanism commanding the emergence of influential cases holds in different scales and trimming is, in some aspects, equivalent to a reduction in scale, where influential values are excluded, new such cases tend to emerge.[9]

As for those accounting variables which may take on negative numbers, figure 8 shows how their distributions may look like so that both positive and negative numbers are proportional to size. Such distributions should be viewed as a juxtaposition of two multiplicative distributions. Accordingly, in order to remove size, two separated ratios should be formed with positive and negative denominators (corresponding to slopes A and $-A$).

Another reason why it is important to assume the correct type of statistical behaviour (additive or multiplicative) of random variables is the fact that formulations adequate to describe inter-relationships amongst variables are distinctly different for each type. Additive formulations implicitly assume that distributions are preserved when variables are added or subtracted, which is not the case of multiplicative data where distributions are preserved when variables are multiplied or divided.

For example, the simplest additive formulation is $x_j = \mu + e_j$, where x_j is explained as

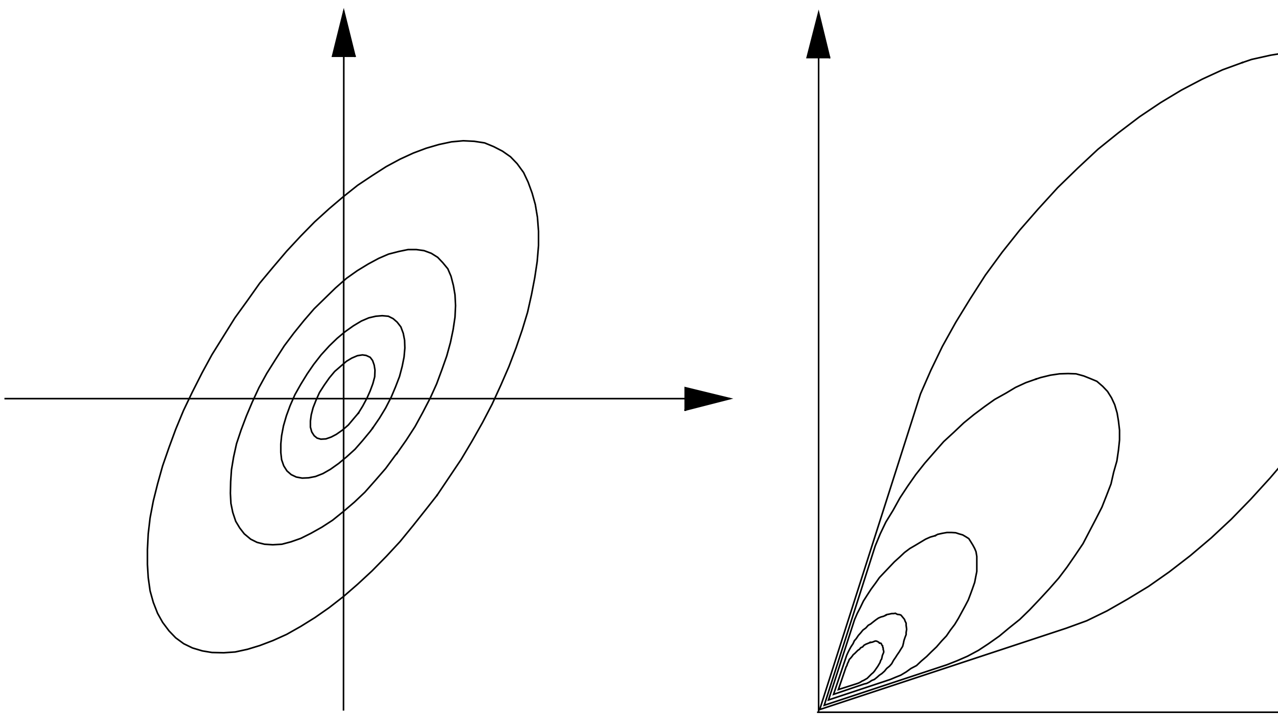

Figure 7: Two stylised representations of bivariate distributions. On the left, a typically additive distribution; on the right, the corresponding multiplicative distribution.

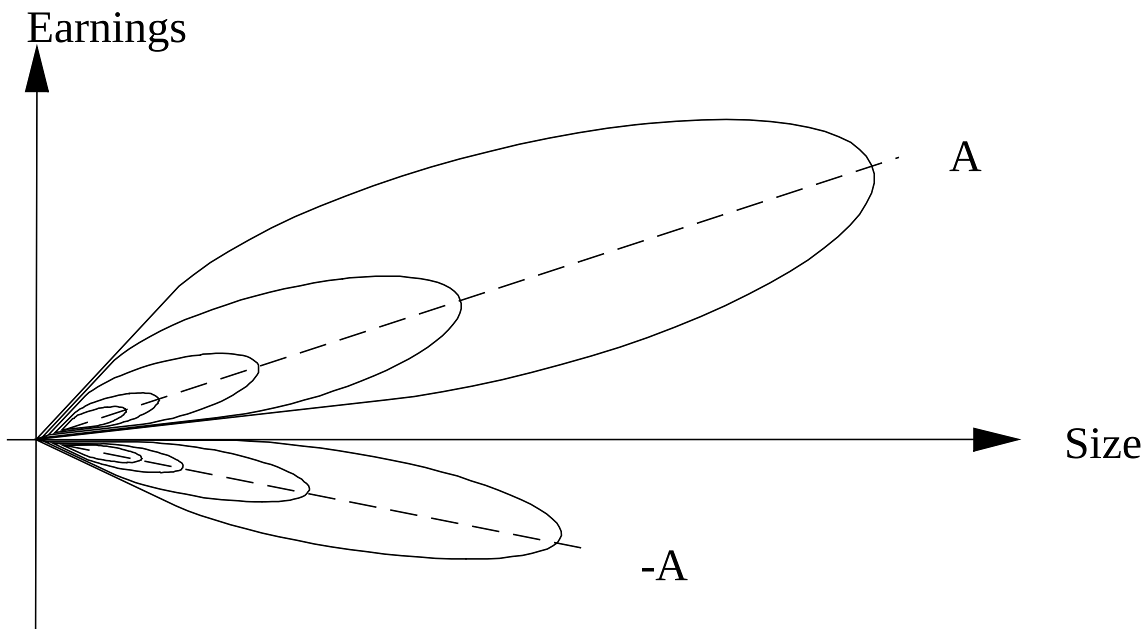

Figure 8: Stylised representation of multiplicative bivariate distribution where one of the variables may take on negative values.

an expected value, μ, plus a random deviation, e_j, specific to j. The multiplicative equivalent would be $x_j = x_0 \, f_j$, where realisation j of x is explained as an expected magnitude, x_0, the same for all cases, *times* a random proportion of x_0, f_j, specific to j.

Where additive variable x is explained, not only as an expected value μ, but also by d, an expected deviation from μ modelling the effect of an extra component of the variance of x, then the above formulation becomes $x_j = \mu + d + e_j$. If d is a component of the variance of two variables, y and x, it is possible to remove it from measurement by assessing y and x in the same subject and then subtracting these two values. For instance, when medical observations are assessed twice in the same patient, *e.g.*, before, (y), and after, (x), a treatment, the difference $y - x$, which measures the effect of the treatment, is free from spurious influences such as those of sex or age because these, being present in both observations, cancel out when subtracted.

The ratio method is the multiplicative equivalent to the above formulation. In the case of accounting data, firm size (the 'spurious' influence to be removed) is a component of the variance of both the numerator, y, and the denominator, x of the ratio. For report j, y and x are described as

$$y_j = y_0 \, e^{s\tau_j} \, f_{yj} \quad \text{and} \quad x_j = x_0 \, e^{s\tau_j} \, f_{xj}$$

where, similarly to (2), y_0, x_0 are expected magnitudes of y, x, $e^{s\tau_j}$ (the same for both variables) is the number of times numbers from report j are expected to be larger or smaller than their expected magnitudes and f_y, f_x are unexplained random proportions.

In the above 'random effects' model,[10] y_0, x_0 are specific to variables (they are independent of the report considered) whereas $e^{s\tau_j}$ measures the relative position of report j irrespective of the variable considered. Since y and x are assessed in the same subject, *i.e.*, in the same report, their ratio will remove $e^{s\tau_j}$, the component of the variance of y and x denoting the effect of report j, thus making reports comparable.

It may be asked whether the two above-mentioned conditions for ratio validity are verified when variables are random. In the following the issue is studied using generative processes obeying the Markov property.[11] Accounting numbers are viewed as the outcome of exponential growth similar to (2) where, according to the random effects model, the continuous rate of growth $s_j = s\tau_j$, reflects the influence of firm size upon the generation of report j irrespective of the variable considered and x_0, y_0 are specific to each variable, reflecting expected scaled size.

Consider y and x, two random variables which are respectively the numerator and the denominator of a ratio. Suppose that, for report j, observed realisations of y and x are generated by exponential diffusion processes

$$d(\log y_j) = s_j dt + \sigma_y dz_{yj} \qquad \text{and} \qquad d(\log x_j) = s_j dt + \sigma_x dz_{xj}, \tag{4}$$

where continuous rates of change $d(\log y_j)$ and $d(\log x_j)$ stem from a deterministic term, $s_j dt$ (the same for both variables and assumed to be constant during generation) plus a random term, σdz_j with $dz_j = z_j\sqrt{dt}$, specific to each variable. z_{yj} and z_{xj} are time-independent standard Normal random variables. The summation of all dt, t, reflects the time during which the generation of report j takes place, typically one year.

By exponentiation, (4) leads to[12]

$$\frac{dy_j}{y_j} = (s_j + \frac{\sigma_y^2}{2})\, dt + \sigma_y dz_{yj} \qquad \text{and} \qquad \frac{dx_j}{x_j} = (s_j + \frac{\sigma_x^2}{2})\, dt + \sigma_x dz_{xj}.$$

Ratios of variables generated as above are expected to be constant, evolving as

$$\frac{y_j}{x_j} = \frac{y_0}{x_0} e^{Z_j},$$

removing s_j, the 'random effect' of size, from measurement. Z_j is a Wiener process[13] with variance $(\sigma_y^2 + \sigma_x^2 - 2\,\rho\,\sigma_y\,\sigma_x)\,t$ (ρ measures correlation between z_y and z_x).

Notice that (4) obeys (3), the robust formulation of scale-invariance. In fact, when modelling *continuous-time* processes, it is the continuous (s) rather than the effective (i) rate of change which is expected to be similar in both components.

The literature on Financial Economics seldom distinguishes between continuous and effective rates because such distinction would be irrelevant given the context. However, when the issue of interest is the existence or not of ratios exhibiting constant expectation, the distinction cannot be overlooked. For instance, if the two processes above were assumed to equate effective rates they should be described as

$$\frac{dy_j}{y_j} = i_j dt + \sigma_y dz_{yj} \qquad \text{and} \qquad \frac{dx_j}{x_j} = i_j dt + \sigma_x dz_{xj}. \tag{5}$$

Even where the deterministic term $i_j dt$ is the same in both components, ratios whose components evolve according to (5) drift exponentially (Tippett, 1990) by $\sigma_x^2/2 - \sigma_y^2/2$ per time unit.

Limitations of deterministic ratios are applicable to the case of random components by assessing their impact on expectation. For example, additive terms δ_y or δ_x in the numerator or in the denominator respectively, are corrected by ratios

$$\frac{y_j - \delta_y}{x_j} \qquad \text{or} \qquad \frac{y_j}{x_j - \delta_x}$$

In addition to such limitations and to those stemming from non-exponential components, the fact that, as seen above, the variance of stochastic ratios is expected to increase with time, should also be considered as a limitation whereby the second condition for ratio validity is contradicted. This issue will be studied elsewhere.

Final Comments

Developments presented in the paper suggest the following: (a) rather than just linearity and convergence, valid ratios require exponential scale-invariance of changes in components; (b) firm size may be viewed as a statistical random effect specific to each report; (c) expected magnitudes of variables may be viewed as size which has been scaled, ratio expectation stemming from such underlying scales; (d) generative processes equating (expected) continuous growth rates proportional to the actual size of the firm lead to ratios which will not necessarily drift.

The paper requires empirical research able to ascertain, namely, the extent to which scale-invariance holds for different ratios and, for individual components, the relative importance of size comparative to other statistical effects. Notice that, since lognormality implies the exponential character of components, the second condition for ratio validity seems to be already well supported by empirical evidence.[14] A question the paper did not pursue is the extent to which deviations from strict lognormality make the measurement dependent on norms or size.

The generative mechanism used in the paper should not be regarded as an attempt to explain how accounting variables are formed in reality.[15] Rather, it is aimed at showing how stochastic components obeying a multiplicative 'random effects' model may lead to ratios exhibiting constant expectation. Drifts, therefore, should no longer be viewed as intrinsic to ratios and valid measurement using ratios is attainable, at least in theory. As a consequence, it makes sense to assess deviations from such valid measurement, or to try to improve ratios so as to bring them closer to such benchmark.

A question the paper may have contributed to answer is whether the use of ratios is motivated by tradition or by well founded considerations. It is now clear that a set of explicit conditions is required for the validity of the ratio method. Moreover, such conditions agree with, and require, the type of data found in accounting reports. Given this, it may then be asked why some authors were led to recommend *ad hoc* approaches to deal with ratios. Reasons seem to be two fold: first, they assumed that financial variables are additive; second, they also avoided discussing the way firm size influences ratio components. Since firm size remained a rather abstract concept, the same applied to the question of whether ratios remove size or not.

Moreover, it seems as though lack of familiarity with numerical methods may have led some authors to make mistakes or, in other cases, to misleading conclusions based on verbatim interpretation of formulæ.

As for mistakes, it is worth exposing the pitfall, coined by Eisenbeis (1977), that 'log-transformed variables give less weight to equal percentage changes in a variable where the values are large than when they are smaller' (p. 877). The pitfall consists of calculating percentages of a log-transformed measurement as this is equivalent to calculating percentages of percentages. Undoubtedly, Eisenbeis' unjust ban on logarithms, later propagated by authors such as Barnes (1982), may have scared away early attempts to fit adequate models to the distribution of ratios while fuelling the use of *ad hoc* techniques such as those proposed by Frecka & Hopwood (1983).

An example of verbatim interpretation of formulæ is the often-quoted statement that almost all of the assumptions required for valid ratio analysis are likely to be violated in practice (Lev & Sunder, 1979). The statement is formally correct, of course, and it might as well be applied to Newton's Laws of Gravity and to many other models considered as good enough approximations in normal circumstances. The statement is misleading because assumptions may be violated without invalidating a methodology. In fact, (a) distortion, in spite of its presence in mathematical models, may be small in specific cases (*e.g.*, the case of non-proportionality); and (b) when weighing accuracy against intuitive interpretation, it may happen that such trade-off is favourable to a less accurate methodology.

In the ratio method this seems to be the case. Ratios, having one degree of freedom only, are able to measure deviations from expected proportions between components. The condition of scale-invariance is a consequence of this: one unique parameter can only deal with the common growth of components. However, by allowing the measurement of percentage deviations from such expectation, ratios go to the point as that is exactly what practitioners need to know in order to make decisions.

Therefore, rather than simply expand the complexity of models so as to cope with limitations of ratios, research on the financial ratio measurement should focus on correcting limitations without changing the specific characteristics of the measurement, a more demanding task.

Notes

1. 'Multiplicative', 'proportionate', 'exponential' or 'logarithmic' are terms variously used in the literature on the growth of firms to designate a family of skewed distributions related to proportionate growth.

2. In scales, original units of measurement are no longer present. A margin of 16%, for example, is a percentage with no units attached. Ratios are scales when components are measured in the same unit, money in this case. When the unit in which the numerator of a ratio is measured is different from that of the denominator, ratios may retain both units as in the case of Earnings per Share, or just one unit, as for Debtors Days.
3. If variables such as Earnings were not related to size, then profitability and dividend yield would be diluted by any increase in size and firms would carefully avoid growing.
4. The relationship between an effective (i) and the underlying continuous (s) rate of change is $s = \log(i+1)$. Continuous rates are also called 'force' of interest, which underlines the continuous character of the influence they model.
5. Infinitesimal changes are those which may be made so small that they approach zero as a limit.
6. For example, an industry-wide average cost of £3,850,000 for food manufacturers in the UK in 1987 would represent only 0.2% of United Biscuits earnings but it would equal or surpass the turnover of 5% of the firms in the same industry.
7. Accumulations of random events tend to be multiplicative, as opposed to additive, because the likelihood of realisations is conditional on the occurrence of a chain of several previous events. Such likelihood thus stems from multiplying, rather than adding probabilities.
8. Lognormality in variables such as Sales, Earnings and Assets, received a great deal of attention in texts on the theory of the growth of firms. In the accounting literature, McLeay (1986*a*) mentions lognormality in variables which are sums of similar transactions with the same sign such as Stocks, Creditors or Current Assets. Recent empirical evidence (Trigueiros, 1995) suggests that lognormality is widespread.
9. A regression where Sales explains Earnings (using data from the UK Electronics industry, 1986) illustrates the above. When the Cook Distance (Cook, 1977) is used to identify influential cases, two firms (G.E. and STC) are singled out as outliers. After trimming these two firms, three new firms (SUNLEIGH, ENGLISH ELECTRIC and BROTHER INTERNAT) become influential. After also excluding the three firms, SYNAPSE COMPUTER emerges with a new Cook Distance of 80, a value which indicates extreme influence.
10. An *effect* models the impact of stratification or co-variance on expectation. Influences such as the two possible sexes are named 'fixed effects' because such stratification is deterministic. Where the stratification is itself a random variable, effects are called random. See, *e.g.*, Snedecor & Cochran (1965, $9^{th.}$ edition, p. 237).
11. The Markov property is verified in those variables where only the most recent realisation contains information useful to predict future realisations. Process (5) or others of the same type are extensively used to model earnings or stock prices as, without the Markov property, investors would be able to avoid risk (Ball & Watts, 1972, pp. 665–666). In the literature on ratios, Lev (1969) is an early example of their use. More recently, Tippett (1990) also used them to induce the behaviour of ratios.
12. Rules of calculus do not apply to stochastic variables. Readers interested in these matters (but with limited mathematical training) may probably do better with texts on Financial Economics (*e.g.*, Dixit & Pindyck, 1994, pp. 59–82) than with those on stochastic processes.
13. Wiener processes (Brownian motion) are continuous-time, obey the Markov condition, have independent increments and are Normal (with zero mean and variance proportional to time).
14. Such evidence is based on those, mostly quoted firms, whose accounts are collected into files such as EXTEL (in the UK). Samples including smaller firms may be not strictly lognormal (Hall, 1987).

15. In the case of variables such as Fixed Assets, Net Worth and the stock underlying Sales, the generative mechanism used in the paper may approach reality or, at least, it is the simplest instance able to do so. However, t in (4) should refer to the birth of the firm, not to the beginning of the year and s_j should not be viewed as constant.

References

Altman, E. (1968). 'Financial Ratios, Discriminant Analysis and the Prediction of Bankruptcy'. *The Journal of Finance*, Vol. 23, No. 4, pp. 589–609.

Ball, R. & Watts, R. (1972). 'Some Time Series Properties of Accounting Income'. *The Journal of Finance*, Vol. 27, No. 3, pp. 663–681.

Barnes, P. (1982). 'Methodological Implications of Non-Normally Distributed Financial Ratios'. *Journal of Business Finance and Accounting*, Vol. 9, No. 1, pp. 51–62.

Beaver, W. (1967). 'Financial Ratios as Predictors of Failure'. *Empirical Research in Accounting: Selected Studies, 1966 Journal of Accounting Research*, pp. 71–111.

Cook, R. (1977). 'Detection of Influential Observations in Regressions'. *Technometrics*, Vol. 19, p. 15–18.

Deakin, E. (1976). 'Distributions of Financial Accounting Ratios: Some Empirical Evidence', *The Accounting Review*, January, pp. 90–96.

Dixit, A. & Pindyck, R. (1994). *Investment Under Uncertainty*. (Princeton University Press).

Eisenbeis, R. (1977). 'Pitfalls in the Application of Discriminant Analysis in Business, Finance and Economics', *The Journal of Finance*, Vol. 32, No. 3, pp. 875-899.

Frecka, T. & Hopwood, W. (1983). 'The Effect of Outliers on the Cross-Sectional Distributional Properties of Financial Ratios', *The Accounting Review*, January, pp. 115–128.

Hall, B. (1987). 'The Relationship Between Firm Size and Firm Growth in the United States Manufacturing Sector'. *Journal of Industrial Economics*, Vol. 35, pp. 583–606.

Horrigan, J. (1965). 'Some Empirical Bases of Financial Ratio Analysis'. *The Accounting Review*, July, pp. 558–568.

Lev, B. (1969). 'Industry Averages as Targets for Financial Ratios'. *Journal of Accounting Research*, Autumn, pp. 290–299.

Lev, B. & Sunder, S. (1979). 'Methodological Issues in the Use of Financial Ratios'. *Journal of Accounting and Economics*, December, pp. 187–210.

McLeay, S. (1986*a*). 'The Ratio of Means, the Mean of Ratios and Other Benchmarks', *Finance, Journal of the French Finance Society*, Vol. 7, No. 1, pp. 75–93.

McLeay, S. (1986*b*). 'Student's *t* and the Distribution of Financial Ratios', *Journal of Business Finance and Accounting*, Vol. 13, No. 2, pp. 209–222.

Snedecor, G. & Cochran, W. (1965). *Statistical Methods*. (Iowa State University Press).

Tippett, M. (1990). 'An Induced Theory of Financial Ratios'. *Accounting and Business Research*, Vol. 21, No. 81, pp. 77–85.

Tippett, M. & Whittington, G. (1995). 'An Empirical Evaluation of an Induced Theory of Financial Ratios'. *Accounting and Business Research*, Vol. 25, No. 99, pp. 208–222.

Trigueiros, D. (1995). 'Accounting Identities and the Distribution of Ratios'. *British Accounting Review*, Vol. 27, No. 2, pp. 109–126.

Whittington, G. (1980). 'Some Basic Properties of Accounting Ratios'. *Journal of Business Finance and Accounting*, Vol. 7, No. 2, pp. 219–232.