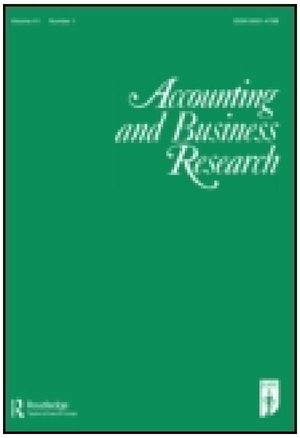


Accounting and Business Research
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/rabr20

Bounded variation and the asymmetric distribution of scaled earnings

Demetris Christodoulou [a] & Stuart McLeay [b]
[a] MEAFA, Faculty of Economics and Business, The University of Sydney, NSW, 2006, Australia E-mail:
[b] University of Sussex
Published online: 04 Jan 2011.

Bounded variation and the asymmetric distribution of scaled earnings

Demetris Christodoulou and Stuart McLeay*

Abstract — This paper proposes a finite limits distribution for scaled accounting earnings. The probability density function of earnings has been the subject of a great deal of attention, indicating an apparent 'observational discontinuity' at zero. Paradoxically, the customary research design used in such studies is built on the implied assumption that the distribution of scaled accounting earnings should approximate a continuous normal variable at the population level. This paper shows that such assumptions may be unfounded, and, using large samples from both the US and the EU, the study provides alternative evidence of a consistently asymmetric frequency of profits and losses. This casts further doubt on the interpretation of the observed discontinuity in the distribution of earnings as prima facie evidence of earnings management. A particular innovation in this paper is to scale the earnings variable by the magnitude of its own components, restricting the standardised range to [–1,1]. Nonparametric descriptions are provided that improve upon the simple histogram, together with non-normal parametric probability estimates that are consistent with the scalar that is proposed. A notable advantage of this approach is that it avoids some of the statistical shortcomings of commonly used scalars, such as influential outliers and infinite variances.

Keywords: accounting earnings; scalars; deflators; boundary conditions; parametric and nonparametric density estimation

1. Introduction

The probability density function of accounting earnings has attracted a great deal of attention, especially in understanding the distribution of earnings in different settings (e.g. Dechow et al., 2003; Burgstahler and Eames, 2003; Leuz et al., 2003; Brown and Caylor, 2005; Peasnell et al., 2005; Yoon, 2005; Daske et al., 2006; Maijoor and Vanstraelen, 2006; Gore et al., 2007). However, a problematic feature of research design in this area is that the use of an earnings deflator may introduce sample bias. Indeed, according to Durtschi and Easton (2005), the observed discontinuity in the distribution of earnings could be a spurious result that is due to scaling by variables that are systematically lower for loss observations than for profit observations. But other researchers (Beaver et al., 2007; Jacob and Jorgensen, 2007; Kerstein and Rai, 2007) have since rejected the arguments put forward by Durtschi and Easton, and claim that the abrupt changes documented at zero earnings are not primarily attributable to scaling.[1] Whilst there is a growing body of work that already questions the assertion that the observed discontinuity is simply a statistical effect, it is noticeable nevertheless that earnings management studies are now being designed with alternative scalars in mind (e.g. Petrovits, 2006; Daniel et al., 2008).

In this paper, we consider another shortcoming that may lead researchers to question the conclusions arrived at regarding the distribution of earnings, and in this context we propose an entirely new scalar with some remarkable properties. In essence, our main concern is that inferences about scaled accounting earnings are generally drawn by presuming normality in the limit. This gives rise to an obvious contradiction in research design when the

* The authors are at the University of Sydney and the University of Sussex. They are grateful for the helpful contributions of Maurice Peat, Maxwell Stevenson and Jo Wells. They also thank the anonymous reviewers for constructive comments.

Further details of the authors' statistical programming code are available at http://meafa.econ.usyd.edu.au/.

Correspondence should be addressed to: Dr Demetris Christodoulou, MEAFA, Faculty of Economics and Business, The University of Sydney, NSW 2006, Australia. E-mail: d.christodoulou@econ.usyd.edu.au.

This paper was accepted for publication in February 2009.

[1] Durtschi and Easton (2005) suggest scaling by the number of shares, and the visual evidence published by these authors suggests that, for earnings per share, the distribution is smoother around zero than was previously documented using other scalars. However, Beaver, et al. (2007) are at odds with this view, and their evidence reveals that the number of outstanding shares tends, in practice, to be higher for losses than profits, which has the effect of shifting scaled loss observations towards zero, and could therefore be responsible for the reduction in the kink in the distribution of earnings when scaled by the number of shares. Jacob and Jorgensen (2007) and Kerstein and Rai (2007) also reject the arguments put forward by Durtschi and Easton, documenting abrupt changes at zero that are not primarily attributable to scaling. The impact of this recent research is now evident in additional testing for such biases – see for example Ballantine et al. (2007) who use a lagged total assets scalar in examining earnings management in hospital trusts, and test accordingly for any systematic difference in the size of the scalar across surplus and deficit trusts.

apparent 'discontinuity' about zero earnings is evaluated in an asymptotic Gaussian setting with a (supposedly) random sampling of firms. In theory, such a design implies that, as the sample size tends towards its population limit, the less the observed 'discontinuity' would be and that eventually it would vanish. If, however, a disproportion is expected around zero regardless of sample size, then expectations should not be based on the normal curve and an alternative framework is required, using either a suitable nonparametric approach or a more appropriate parametric model. This paper introduces such a framework, together with the appropriate scalar.

To illustrate, Figure 1 provides a histogram of the central empirical frequencies of net income scaled by sales, using the sample that is employed in this research study. The lack of fit to the normal curve that can be seen around zero is generally interpreted as prima facie evidence of earnings management, but, as mentioned already, this inference depends on the appropriateness of the normality assumption for the population.[2] Figure 1 also clearly demonstrates how inferences concerning the conjectured discontinuity depend on the origin and bin width of the histogram estimator. In this respect, it is usual to examine the difference between the observed probability of earnings for the i^{th} bin next to zero (estimated as $p_i = n_i/n$, where n is the total number of observations and n_i the number of observations that fall in the i^{th} bin) and an expected probability that is calculated as the average of the two adjacent bins, i.e. $E(p_i) = (p_{i-1} + p_{i+1})/2$. If the normalised difference $|p_i - E(p_i)|/std.error$ is large, then the observational disproportion around zero is taken as an indication of earnings management. However, the histogram itself defines the weight of the observed probabilities p_i, and Figure 1 shows how a bin width and origin that are selected to separate the groupings at zero will bias the nonparametric representation in a way that emphasises this disproportion.[3]

In view of the implicit shortcomings outlined above, and in the light of the debate initiated by Durtschi and Easton (2005), we motivate this paper by first explaining why we might expect an asymmetric shape in net income, not simply as the outcome of earnings management but for more fundamental reasons. Then we describe an approach to modelling earnings using a scalar which reflects the magnitude of the components of income, from which we derive a measure of scaled accounting earnings with known bounds. A major benefit of the known range of variation under the transformation is that the distributional space remains standardised for all observations regardless of the sample size and the degree of heterogeneity. Indeed, we are able to show how the variation of scaled accounting earnings asymptotically approximates the limits that are reached in the extreme cases where firms report either zero costs or zero sales. To examine the shape of the earnings distribution, a kernel density estimator is employed that provides a more detailed and unbiased description than the commonly-used histogram estimator, showing that net income is consistently asymmetric across samples drawn from different economic regions and different accounting jurisdictions. The paper also derives a bounded parametric density function with the ability to accommodate such asymmetry, which is analytically superior to the standard normal. Given the above, we conclude with evidence that the key characteristic of scaled accounting earnings is its asymmetry, and it is suggested in the final discussion that this may be attributable to a great extent to firm-level heterogeneity effects, as the asymmetry is removed when we examine mean-adjusted densities at the firm level.

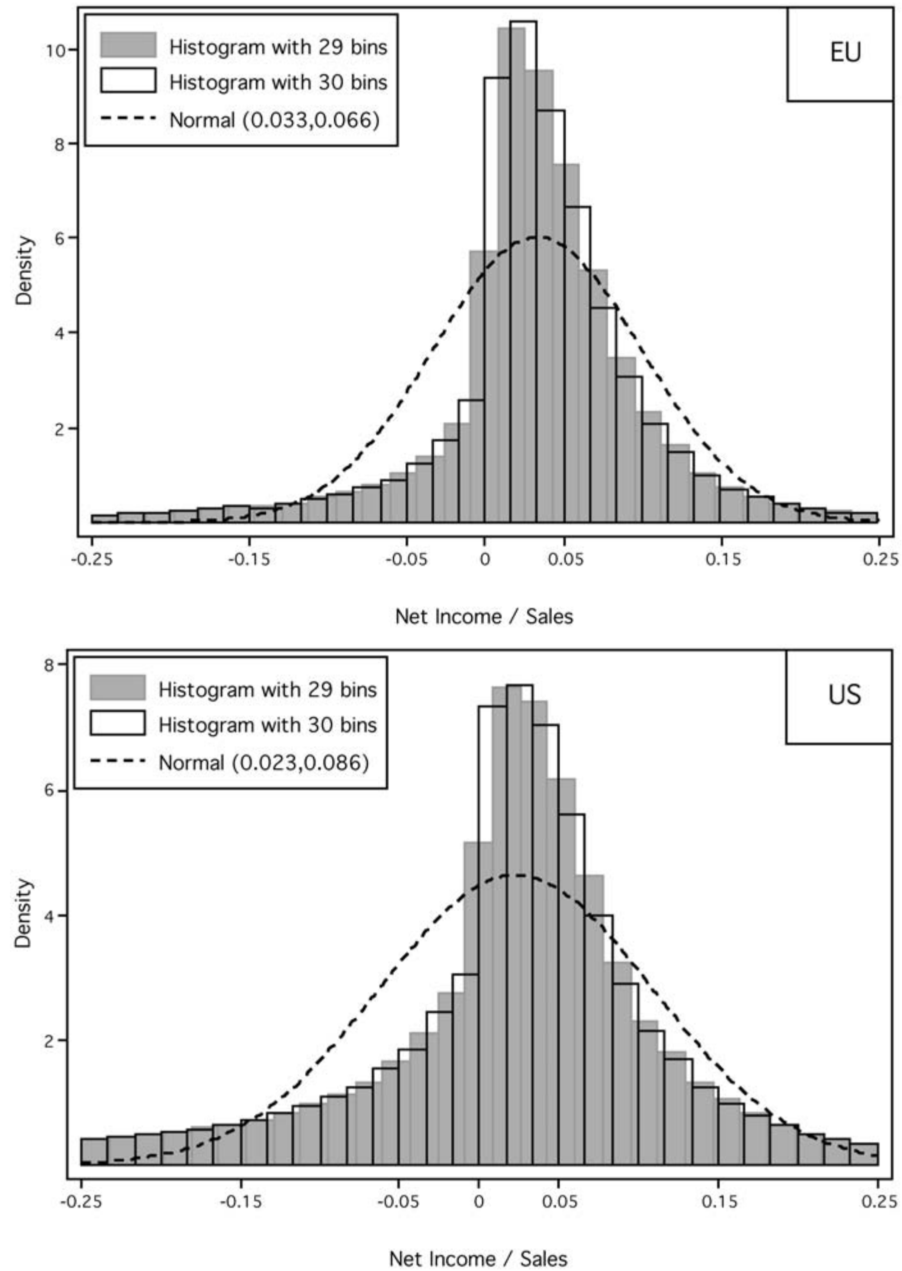
2. The distribution of accounting earnings

In this paper, we argue that normality in scaled earnings is not consistent with the character of the accounting variables involved in calculating and

[2] This type of analysis dates back to Burgstahler and Dichev (1997) and Degeorge et al. (1999), who investigate bottom-line net income, as we do in this paper. They scale net income by market value and number of shares respectively. For a sample of firm-year earnings, they each observe a disproportionate frequency around zero and attribute this result to the upwards management of earnings, in the sense that listed firms with small losses attempt to beat the benchmark of zero and thus report small profits. Degeorge et al. (1999) test the null hypothesis that the density of earnings is smooth at the point of interest, assuming that the change in probability from bin i to bin i+1 is approximately normal. We infer that this implies normality in the parent distribution as the sum of normal variates is itself normal.

[3] It is a common practice to round the number of bins to the nearest even integer and then to compute the expected value $E(p_i)$, so that the bins are separated at zero. Figure 1 illustrates the potential misrepresentation that is inherent in this approach by comparing two histograms with the same origin and range, with one separating at zero and the other having an odd number of bins and therefore not separating at zero. A histogram with 30 bins generates the well-documented difference in probabilities around zero, with $p_i - E(p_i)$ in the negative bin adjacent to zero equal to −0.0497 for the EU and −0.0305 for the US, the difference in the positive bin adjacent to zero being equal to 0.0467 for the EU and 0.0325 for the US. However, when we estimate a histogram with 29 bins, we find little difference between the observed and expected probabilities for the bin which contains the point zero (–0.0097 for the EU and −0.0007 for the US), and what is more, the shape no longer implies an observational disproportion at zero.

Figure 1
The frequency distribution of net income scaled by sales



Note: The histograms describe the frequencies of net income scaled by sales within the range [–0.25, 0.25], for 48,563 EU firm-years and 104,170 US firm-years covering the period 1985–2004. The shaded histogram has 29 bins and the histogram drawn in outline has 30 bins. In the latter case, bin separation at zero conveys the impression of a 'discontinuity' at zero, whereas the empirical frequencies depicted by the shaded histogram appear much smoother. For each plot, the superimposed dashed curve represents the normal curve fitted with the mean and standard deviation of the sample, displayed here over the range [–0.25,0.25].

scaling earnings. To start with, the double-entry bookkeeping system that generates earnings requires a one-to-one correspondence between debits and credits that cannot be related to the randomness of normal probability laws (Ellerman, 1985; Cooke and Tippett, 2000). Indeed, the seminal work of Willett (1991) on the stochastic nature of accounting calculations provides a general proof that bookkeeping figures, and earnings in particular, may be better represented as deterministic measures of random variables. Given that double-entry bookkeeping is a series of algebraic operations on ordered pairs of such numbers, it is evident that the accounting process creates an endogenous matrix of linearly related information, which will include all of the components of earnings and each of the book-based scalars commonly used in accounting research. These relationships will in turn determine the variance of scaled earnings. As shown elsewhere, earnings scaled by assets – although clearly non-normal – will most likely have finite variance; on the other hand, earnings scaled by sales will tend towards infinite variance; and earnings scaled by equity will be Cauchy, with infinite variance and no location (McLeay, 1986).[4]

As for the observed disproportion in scaled earnings around zero, this may be the result of a number of other factors. First, it is now generally accepted that the asymmetry is attributable in part to transitory components, which tend to be larger and more frequent for losses than for profits, and with differing implications for corporate income taxes (Beaver, McNichols and Nelson, 2007). Indeed, as these authors argue, while income taxes will tend to push profits towards zero, transitory items will tend to pull loss observations away from zero. Therefore, it would be reasonable for us to infer that the great concentration in small profits and the asymmetry around zero might result as much from fiscally-driven downward pressure on reported profits as it does from upward pressure to avoid reporting losses. Age, size and listing requirements also offer themselves as partial explanations for asymmetry about zero. Undoubtedly, age is linked to size, the latter usually being proxied by sales or total assets, each of which is known to grow exponentially. Larger size firms are shown to exhibit more stable income streams and an accelerated mean-reversion following a loss (e.g. Prais, 1976), especially following extreme negative changes (Fama and French, 2000). Moreover, listing requirements favour profitable firms, as candidates for listing are required to show evidence of generating sustainable profits. Hence, as markets grow, the number of newly quoted firms increases, and these relatively small firms are most likely to report small profits in their first years of listing.[5]

The likely disproportion around zero income has also been attributed to risk averse behaviour. Kahneman and Tversky (1979), on the psychology of risk aversion, refer to a reflection effect that is concave for profits and convex for losses, and steeper for losses than for gains, yielding an S-shaped asymmetric function about zero income.[6] Along similar lines, Ijiri (1965) characterises the zero point in accounting earnings as the modulator of asymmetry, imposed on the firm either externally or internally. However, in addition to these theoretically-grounded explanations, where zero in earnings acts as a threshold, we should also recognise that the observation of a disproportion around zero in a sample of company earnings could arise in a variety of statistical contexts, of which we comment here only on the most important. First, a gap in observational frequency may be the result of incomplete sampling, which may cause the low density below zero. However, this seems not to be the case, as the samples that are selected are generally consistent and as large as possible across years and firms. Also, it may be supposed that the sample contains observations from more than one population. Such mixtures of samples essentially imply a distribution derived from distinct populations with dissimilar moments, yet this also seems not to be likely, as the firms that are pooled tend to operate under shared economic conditions with respect to competitiveness and maximisation of stakeholders' wealth. Although local modes might be noticeable amongst pooled losses on the one hand and pooled profits on the other hand, plainly it would be inappropriate to classify a firm that may report a loss one year and a profit the next as either a loss-making entity or a profit-making entity.

In our opinion, a likely analytical explanation of the problem is that the non-normal shape of the curve arises directly from the theoretical funda-

[4] Note that, if we scale one pure accounting variable by another, the mathematical boundaries imposed by the double-entry system on the scaled variable are predictable, including a lower bound of zero for total sales over total assets (Trigueiros, 1995), bounds of [0,1] for current assets over total assets (McLeay, 1997), etc. An attempt at describing the complete set of scaled accounting variables is included in McLeay and Trigueiros (2002). There is further discussion and evidence regarding the dynamics of scaling by geometric accounting variables in Tippett (1990), Tippett and Whittington (1995), Whittington and Tippett (1999), Ioannides et al. (2003), Peel et al. (2004), and McLeay and Stevenson (2009). This is particularly relevant to the use of scaled accounting variables in panel analysis. The scaling issue has been addressed also in the context of equity valuation modelling – see Ataullah et al. (2009) for a recent discussion.

[5] To demonstrate this point, Dechow et al. (2003) compare a sample of firms that report earnings in the vicinity of zero and find more small profits in firms that have been listed for two years or less, and that those firms reporting small losses appear significantly larger in size than those reporting small profits.

[6] Kahneman and Tversky (1979) propose an S-shaped reflection effect that is concave for profits and convex for losses on the basis that 'the aggravation that one experiences in losing a sum of money appears to be greater than the pleasure associated with gaining the same amount' (p. 279).

mentals of the population.[7] This consideration is deeply rooted in statistical modelling, and it is the type of behaviour that we suspect we are dealing with here. In particular, if the earnings probability density function (PDF) is part of the non-normal class of densities, then the asymptotic Gaussian assumptions that are commonplace in earnings research are insufficient.

3. A model of scaled earnings with bounded variation

Consider the non-negatively distributed integers $\{x, y, z \geq 0\}$ with a distribution of unknown character and let these be associated as follows:

$$z_{it} = x_{it} - y_{it} \tag{1}$$

with i=1,2,...,I and t=1,2,...,T, so that it indicates an observation from an N=$I \times T$ sample. Now, to standardise with bounded variation, divide by $x_{it} + y_{it}$:

$$\frac{z_{it}}{x_{it} + y_{it}} = \frac{x_{it} - y_{it}}{x_{it} + y_{it}} \tag{2}$$

where $x_{it} + y_{it} \neq 0$ and $[x_{it} - y_{it}] \leq [x_{it} + y_{it}]$. By separating the right-hand side of Equation (2) into two distinct fractions as follows:

$$\left[\frac{z_{it}}{x_{it} + y_{it}}\right]_{-1}^{1} = \left[\frac{x_{it}}{x_{it} + y_{it}}\right]_{0}^{1} - \left[\frac{y_{it}}{x_{it} + y_{it}}\right]_{0}^{1} \tag{3}$$

it is evident that the [–1,1] boundary conditions on the left-hand side are induced because x_{it} and y_{it} are components of the common denominator $x_{it} + y_{it}$. It follows that the standardised sample space [–1,1] is defined by the difference between two [0,1] integrals.

Now consider the general case for any firm, in any accounting period, where expenditure is incurred in the process of generating revenues, resulting in its most basic form in the following accounting identity:[8]

$$\text{Earnings} \equiv \text{Sales} - \text{Costs} \tag{4}$$

At the primary level of aggregation, each of these two variables is a non-negative economic magnitude. This statement may at first seem counter-intuitive in the context of the double entry system, but is clearly evident in the negative operator in Equation (4). The sign is incorporated as an exogenous constraint, and as a result the Earnings variable can take any value on the real number line (i.e. as either profits or losses). In this paper, we exploit this natural positive variability in accounting aggregates in order to derive a scaled form of earnings. Applying the model described in Equation (2) to the difference between Sales and Costs, and deflating by the total magnitude of the two, we derive the variable of interest in this paper, Scaled Earnings E', as follows:

$$E' = \frac{\text{Earnings}}{\text{Sales} + \text{Costs}} = \frac{\text{Sales} - \text{Costs}}{\text{Sales} + \text{Costs}} \tag{5}$$

By giving mathematical support to the range of variability in this way, Equation (5) transforms Earnings into a measure of proportionate variation.[9] That is to say, at the limit, when Sales (Costs) equal zero, then it follows that E' will be equal to minus (plus) one. Over this range, E' will be distributed in the following manner:

$$E' = \frac{\text{Sales} - \text{Costs}}{\text{Sales} + \text{Costs}} \begin{cases} -1 \text{ when Sales} = 0 \\ < 0 \text{ when Sales} < \text{Costs} \\ = 0 \text{ when Sales} = \text{Costs} \\ > 0 \text{ when Sales} > \text{Costs} \\ +1 \text{ when Costs} = 0 \end{cases} \tag{6}$$

Thus, with profitable returns on a scale above 0% to 100%, and negative returns below 0% to -100%, Scaled Earnings E' can be interpreted as a percentage return on the total operating size of the firm, where size is measured in terms of the magnitude of all operating transactions that take place within a financial year.[10] The effect of scale will appropri-

[7] See Cobb et al. (1983) for the statistical justification for treating non-normality as the general case, and symmetry as a special case.

[8] Here, the analysis is deliberately simple, although it is well known that Earnings may be described as a more complex summation. For example, it is the case that most firms also generate other types of revenue in addition to their Sales. Yet the same result applies if, for instance, Earnings were to be defined more comprehensively as the difference between all revenues and all expenditures, rather than just between Sales and Costs.

[9] As discussed earlier, it is usual in financial analysis to scale measures of profitability and performance by size variables, such as the number of outstanding shares, the market value or the beginning-of-the-year total assets, which tend to be selected in an ad hoc manner. Scaling in this way is intended to deal with issues arising from sample heterogeneity, mostly resulting through composite size biases, while the use of a lagged size measure can help to mitigate the autocorrelation problem. A sensitivity analysis on the sample employed in this study verifies that the deflator Sales+Costs is highly correlated with other commonly used size measures, including the market value, a figure which is not taken from the financial statements (for our sample, the coefficient of correlation between Sales +Costs and market value is 0.8303).

[10] As E' is the index of two variables that are measured in terms of the same numeraire, the scaled earnings variable that is proposed here is a numeraire-independent quantity. Another interesting property is the one-to-one correspondence between the components of Scaled Earnings. By denoting E'=Earnings/(Sales+Costs), S'=Sales/(Sales+Costs) and C'=Costs/(Sales +Costs), and recognising that S'+C'=1, then by applying expectations operators E(.), the following theoretical properties can be seen to hold: expected means $\mu_{E'} = \mu_{C'} - \mu_{S'}$; standard deviations $\sigma_{E'} = 2\sigma_{S'} = 2\sigma_{C'}$; skewness $\beta_{1E'} = \beta_{1S'} = -\beta_{1C'}$;

ately increase as Sales and/or Costs deviate from zero, dramatically shrinking the tails of earnings without the need to eliminate extreme observations. Furthermore, with known boundaries, the symmetry of E' is no longer a desired property. Trigueiros (1995) generalises this argument by showing that, when any accounting aggregate following a process of exponential growth is divided by another, the scaled variable will be characterised not by symmetry but by skewness.[11]

As a final point, it should be considered here how scaling by Sales plus Costs might compare with scaling by Sales alone. Formally, it is the case that $-1 \leq$ (Sales–Costs)/(Sales+Costs) ≤ 1 whilst $-\infty <$ (Sales–Costs)/Sales ≤ 1, and the relationship between these two measures is distinctly nonlinear – a concave function with no point of inflection. Indeed, the rate of change in the two measures is similar at one point only, when Costs are 41% of Sales.[12] Furthermore, there are only two points at which the two measures give an identical result, i.e. at breakeven when Sales and Costs are equal, where (Sales–Costs)/(Sales+Costs) = (Sales–Costs)/Sales = 0, and in the extreme case of zero Costs, where (Sales–Costs)/(Sales+Costs) = (Sales–Costs)/Sales = 1. Of course, when Sales = Costs = 0, the company will have ceased operating. As mentioned earlier, the standardised range of E' is a particularly useful property of the Scaled Earnings variable, whilst scaling by Sales alone results in an infinite left-hand tail and generates outliers accordingly. Moreover, as a scalar, Sales fails the Durtschi-Easton test, being systematically lower for loss observations than it is for profit observations. In other words, losses tend to be associated with lower outputs than expected. In contrast, the scalar proposed here, Sales+Costs, corrects for this bias because losses are attributable not only to falling Sales but also to increasing Costs. That is, whilst Sales+Costs = 2 × Sales = 2 × Costs at breakeven point, Sales+Costs is greater than 2 × Sales (but lower than 2 × Costs) when there is a loss, and lower than 2 × Sales (but greater than 2 × Costs) when there is a profit.

4. A generalised probability function for scaled earnings

It is argued above that, given the bounded character of profitability, and the expectation of population asymmetry about zero, the normal distribution is inappropriate for describing Scaled Earnings E'. Nevertheless, it is possible to express the standard normal integral $z \sim N(0,1)$ as a function $g(.)$ of the unknown distribution of E' conditional on a set of parameters ω, so that $z = g(E'|\omega)$. Following Johnson (1949), E' may be expressed as a linear approximation to the standard normal z, conditional on $\omega = \{\xi, \lambda, \gamma, \delta\}$ with location ξ, scale λ and shape parameters γ and δ, as follows:

$$z = \gamma + \delta\, f\left(\frac{E' - \xi}{\lambda}\right), \text{ where } \delta, \lambda > 0. \tag{7}$$

Recalling that the boundary conditions for the Scaled Earnings variable E' are known to be [–1,1], it is evident that, as E' shifts location from the lower bound $\xi=-1$ to the upper bound $\xi+\lambda=1$, resulting in scale $\lambda=2$, $\left(E'+1\right)/2$ will relocate from 0 to 1, with γ and δ giving shape to the distribution. It follows that a suitable translation of $f(.)$ in Equation (7) is the logit function

$$\ln\left(\left(E' - \xi\right)/\left(\xi + \lambda - E'\right)\right) = \ln\left(\left(E' + 1\right)/\left(1 - E'\right)\right)$$

which increases monotonically from $-\infty$ to ∞ as $\left(E'+1\right)/2$ increases from 0 to 1. Thus, the bounded function for E' belongs to the particular class of Johnson bounded distributions that are described in Equation (8) (see box below), where $\gamma + \delta \ln\left(\left(E'+1\right)/\left(1-E'\right)\right) = z \sim N(0,1)$ is now a reasonable approximation to the standard normal conditional on the estimation of γ and δ, and

$$\Phi\left(E' | \xi = -1, \lambda = 2, \gamma, \delta\right) = \left\{\frac{\delta}{\sqrt{2\pi}} \frac{2}{(E'+1)(1-E')} \exp\left\{-\frac{1}{2}\left(\gamma + \delta \ln\left(\frac{E'+1}{1-E'}\right)\right)^2\right\}\right. \tag{8}$$

kurtosis $\beta_{2E'} = \beta_{2S'} = \beta_{2C'}$; and product-moment correlations $\rho_{E'S'} = -\rho_{E'C'} = -\rho_{C'S'} = 1$. Note that these results may also be obtained from the identities: $S'/(S'+C') = 1 - C'/(S'+C')$ and $(S'-C')/(S'+C') = 2S'/(S'+C') - 1 = 2C'/(S'+C')$.

[11] The proponents of multiplicativity in accounting variables include, amongst others, Ijiri and Simon (1977), Trigueiros (1997) and Ashton et al. (2004). This notion is commonplace in the industrial economics literature, where firm size (generally proxied by sales) is treated as a stochastic phenomenon arising from the accumulation of successive events (see the influential work of Singh and Whittington, 1968).

[12] Applying the quotient rule $u/v' = (u'v - uv')/v^2$ to the differentiation of (Sales–Costs)/(Sales+Costs) with respect to Sales, so that u = Sales–Costs with $u' = 1$ and v = Sales+Costs with $v' = 1$, the first derivative is equal to $(v-u)/v^2 = 2 \times$ Costs/(Sales+Costs)^2. Similarly, differentiating (Sales–Costs)/Sales with respect to Sales gives Costs/Sales^2. To obtain the unique point where the two functions have the same sensitivity to Sales, we set the two derivatives equal, and find that Costs = $(\sqrt{2}-1) \times$ Sales $\approx 0.41 \times$ Sales. That is to say, the rate of change in the two functions is equal at the point where Costs is equal to 41% of Sales. We are grateful to Jo Wells for suggesting this solution.

where E' may take values within the specified range but not the extreme values -1 and 1. It is important to note that Equation (8) describes distributions with high contact at both ends, and guarantees the finiteness of all moments. Furthermore, it has the ability to accommodate frequencies that may take any sign of skewness and any value of kurtosis, which may also be decentralised or bimodal. Another point of interest is that, if E' is distributed in a symmetrical manner, then Equation (8) will not alter the symmetry, whereas other types of transformation usually have an adverse effect.[13]

As for recovering the parameters of Equation (8), it is known that maximum likelihood estimation is oversensitive to large values of higher-order moments when there is high concentration around the mean (Kottegoda, 1987). As a preliminary investigation of the shape of Scaled Earnings suggests that we must expect particularly concentrated peakedness, we use here the more flexible fitting method of least squares, originally developed by Swain et al. (1988), which is known to yield significantly improved fits (Siekierski, 1992; Zhou and McTague, 1996). Additionally, it provides for the option to estimate with even narrower limits for $\xi > -1$ and/or $\xi + \lambda < 1$, if the empirical data suggest this. Details of the proposed steps for recovering the parameters are provided in the Appendix.

5. Examining the shape of scaled earnings

In this paper, in addition to the histogram, we also employ a more flexible nonparametric tool – the kernel density estimator – in order to obtain a smoothed representation of the shape of Scaled Earnings E'. As this appears to be a new approach in the context of earnings analysis, a brief overview is provided here in order to set out the main advantages over the histogram.

Kernel density estimation arranges the ranked observations into groups of data points in order to form a sequence of overlapping ‘neighbourhoods’ covering the entire range of observed values. Each localised neighbourhood is defined by its own focal mid-point m, and the number of data points in any neighbourhood depends on the selection of a bandwidth b. Thus, for the continuous random variable E' with independent and identically distributed (IID) observations, the kernels constituting the estimator are smooth, continuous functions of the overlapping neighbourhoods of observed data, and the kernel density estimator is defined as a summation of weighted neighbourhood functions as follows:

$$\hat{f}\left(E'\right) = \frac{1}{Nb}\sum_{it=1}^{N} K\left(\frac{E'_{it} - E'_{m}}{b}\right) \tag{9}$$

for it=1,2, . . . ,N firm-year observations, m=1,2, . . . <N mid-points of neighbourhoods with bandwidth b, and a kernel density function K that integrates to one.[14]

It can be shown that the histogram estimator is a limiting case of Equation (9). The histogram is a function of a fixed number of non-overlapping bins, and lacks flexibility by comparison with the kernel estimator as an equally weighted kernel function K=1 is implied for all observations. The resulting estimation is neither smooth nor continuous, and, as demonstrated earlier in Figure 1, the histogram can be particularly misleading for frequencies with significant localised variability.

For the more flexible kernel density estimator, we are faced with the following trade-off: the wider the bandwidth b, the smaller the number of estimates of K. Since the range of E' is standardised, the neighbourhood mid-points m are spaced from -1 to 1. At the limit, for b=2, only one symmetrical kernel about zero is implied, while as $b\rightarrow 0$ the number of kernels increases. The selection of bandwidth is of critical importance therefore, as it defines the number of observations required for estimation with respect to each focal mid-point. Since the ultimate aim here is to examine localised variability surrounding zero earnings, our choice of the Parzen kernel function K together with Silverman’s rule of thumb regarding bandwidth b provides the level of detail in representation that is required.[15]

[13] Previous studies have also employed generalised non-normal distributions for describing frequencies of scaled accounting variables (ratios). For example, Lau et al. (1995) suggest the Beta family and the Ramberg-Schmeiser curves; and Frecka and Hopwood (1983) the Gamma family of distributions. By comparison to the Johnson, these are subordinate bounded systems that can only handle a limited shape of curves. Note also that the non-existence of moments in the distributions of scaled accounting variables causes severe problems when the transformed variable is used in multivariate statistical analysis (Ashton et al., 2004).

[14] The kernel density estimator in Equation (9) was originally proposed by Rosenblatt (1956). For a further description and a detailed bibliography of nonparametric density estimation, see Härdle (1990) and Pagan and Ullah (1999); on the choice between kernels K, see Müller (1984).

[15] For $\zeta_{it} = (E'_{it} - E'_{m})/b$, the Parzen (1962) kernel weighting is as follows:

$$K[\zeta_{it}] = \left\{\left(4/3 - 8\zeta_{it}^2 + 8|\zeta_{it}|^3\right)\forall|\zeta_{it}| \leq 1/2\right\};$$

$$\left\{\left(8(1-|\zeta_{it}|)^3/3\right)\forall 1/2 < |\zeta_{it}| \leq 1\right\};\left\{0\forall|\zeta_{it}| > 1\right\}.$$

Silverman’s rule of thumb requires that the mean squared error is minimised during the selection of bandwidth $b,$ under the

6. Analysis

The sample consists of listed companies in the EU and the US, covering the time period 1985–2004. We include all firms listed during this period, whether active or inactive at the census date, and to ensure comparability with previous studies such as Durtschi and Easton (2005) and Beaver, McNichols and Nelson (2007), we eliminate all financial, utility and highly regulated firms (SIC codes between 4400–4999 and 6000–6999), leaving 54,418 usable firm-year observations for the EU and 140,209 for the US.[16] This selection of non-financial and non-utility firms is also appropriate given the operating orientation of the Scaled Earnings expression in Equation (5). The financial statement data for the EU has been collected from Extel, and the calculation of Scaled Earnings is based on Extel items EX.NetIncome (after tax, extraordinary and unusual items) and EX.Sales, with Costs = EX.Sales – EX. NetIncome. The US data is taken from Compustat, using item 12 for Sales and item 172 for Net Income, again with Costs as the difference between the two.

With regard to extreme observations, as explained above, we consider these to be characteristic of accounting data. Yet, in financial research, it is commonplace to remove such observations as outliers even though, paradoxically, the hypothesised distribution is often assumed to have infinite tails.[17] In contrast, a density with known support, such as that of Scaled Earnings E', does not justify the elimination of data that lie close to the tail-end. For E', the true extremities lie *exactly* at the limits of the function, that is, at $E'=-1$ where Sales=0 and at $E'=1$ where Costs=0. Figure 2 shows the extent of the concentration of the pooled dataset at these limits, highlighting the observations with zero Sales on the left (807 for the EU and 4,799 for the US) and those with zero Costs on the right (524 for the EU and 1,176 for the US).[18] For further analysis, we exclude these observations as they represent firm-years with truly extreme reporting behaviour. Indeed, with regard to parametric density fitting, they represent data points that make no contribution to the surrounding local variability and therefore are an artificial source of multi-modality.

The final working sample comprises 53,087 firm-year observations for the EU and 134,234 for the US. Table 1 provides summary statistics, both for losses and for profits. In each location, the standard deviation of losses (0.2308 in the EU and 0.2772 in the US) can be seen to be far greater than that of profits (0.0521 in the EU and 0.0651 in the US). Figure 2 helps us to understand why it is that losses are more variable than profits. The bounded transformation of Net Income into Scaled Earnings reveals that losses are inclined to populate their entire permissible region, while this is not the case with respect to profits. Indeed, we find that there is only a very small likelihood that E' might exceed 0.5, at which point Sales would be more than three times larger than Costs. This asymmetry in the tails is reflected in the distribution of losses and profits around zero, as shown in Figure 2 by the frequencies in the central percentile. By looking more closely in the vicinity of zero in this way, it is clear that what has been characterised previously as a shortfall in small loss observations appears to be attributable to asymmetry defined by point zero, which is consistent with the fact that the tail densities are much greater for losses than for profits. This asymmetric tendency is further reflected in the medians for losses and for profits reported in Table 1 (EU median loss -0.0433, median profit 0.0223; US median loss -0.1009, median profit 0.0264).

Table 1 also gives a breakdown of the EU sample by member state, based on the location in which each of the firms is domiciled. In most of the smaller sub-samples (Austria, Finland, Greece, Ireland, Luxembourg and Portugal), we find that the minimum and/or the maximum of observed Scaled Earnings is far from the respective sample limit of either -0.9999 or 0.9999 (i.e. excluding zero Sales and zero Costs). In the larger jurisdictions, however,

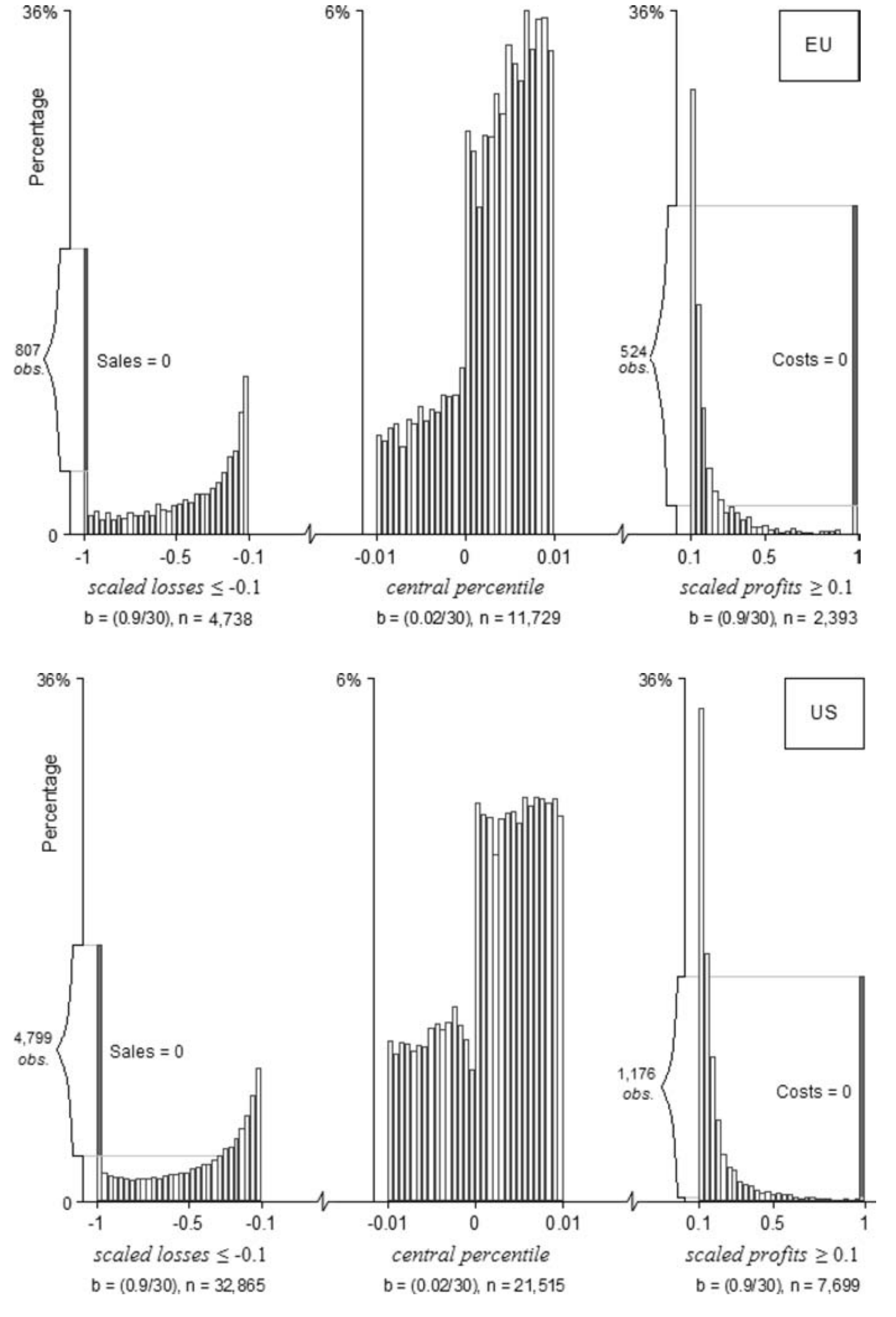
criterion $b = (0.9/N^{1/5}) \times \min\{\sigma, IQR/1.349\}$ where σ is the sample standard deviation and IQR the inter-quartile range (Silverman, 1986; Salgado-Ugarte et al., 1995).

[16] We also excluded a number of observations relating to German firms that habitually reported zero Net Income by adjusting their depreciation expenses and reserves accordingly. In this respect, 127 firm-year observations relating to 32 firms domiciled in Germany were deemed not to be usable. Degeorge et al. (1999) point out that break-even firms such as these will emphasise the discontinuity at zero, as scaling disperses non-zero earnings observations but not those which are exactly zero.

[17] It is shown elsewhere, in McLeay and Trigueiros (2002) and Easton and Sommers (2003), for instance, that the deflation of earnings commonly leads to a number of extreme observations which, once removed, give way to other observations that take their place as new outliers. Indeed, Easton and Sommers (2003) examine the multivariate distribution of market capitalisation, book value and net income and find that up to 25% of the sampling distribution would have to be removed to deal with the statistical problems arising from outliers.

[18] The number of firms reporting zero Costs at least once is as follows: EU 212, US 797; more than once: EU 102, US 207. Those reporting zero Sales at least once is: EU 331, US 1,785; more than once: EU 170, US 1,019. The reporting of zero Costs appears to be unrelated to corporate domicile or industry, whereas this is not the case for zero Sales firm-year observations (i.e. in mineral, oil and gas extraction [SIC 1000–1499]: EU 385, US 1141; in pharmaceutical preparations, medicinal and biological products [SIC 2830–2839]: EU 69, US 856).

Figure 2
Limits to the distribution of scaled earnings



Note: The *y*-axis labels indicate the number of observations at the limits of the function (i.e. EU zero Sales 807, EU zero Costs 524; US zero Sales 4,799, US zero Costs 1,176). Each histogram is plotted from n firm-year observations, with bin width b and 30 bins. The origin for each histogram is set to the left-hand limit.

where markets are deeper and data are not sparse, the frequencies of Scaled Earnings tend to cover the full range. It can be seen from Table 1 that the skew estimate is consistently negative in the larger jurisdictions, at levels that reflect the estimate of −3.6691 for the EU sample as a whole. Finally, the kurtosis of the observed frequencies is high in all sub-samples, reflecting not only the concentration around the mean but also the finiteness of tails, particularly for profits (see Balanda and MacGillivray, 1988). Allowing for the distorting effect of small sub-sample size on some estimates,

Table 1
Final sample and summary statistics

Pooled for 1985–2004	*Observations*			*Summary statistics for scaled earnings*						
	Losses	*Profits*	*Total*	*Minimum*	*Mean*	*Median*	*Maximum*	*Standard deviation*	*Skew*	*Kurtosis*
EU	12,138	40,949	53,087	–0.9999	–0.0077	0.0152	0.9999	0.1420	–3.6691	24.4571
EU Profits			40,949	0.0005	0.03404	0.0223	0.9999	0.0521		
EU Losses			12,138	–0.9999	–0.1486	–0.0433	–0.0000	0.2308		
US	55,931	78,303	134,234	–0.9998	–0.0712	0.0074	0.9999	0.2299	–2.1689	8.3259
US Profits			78,303	0.0000	0.0433	0.0264	0.9999	0.0651		
US Losses			55,931	–0.9998	–0.2315	–0.1009	–0.0000	0.2772		
EU, by member state										
Austria	192	756	948	–0.7815	0.0184	0.0112	0.9015	0.1161	2.8207	33.5815
Belgium	250	914	1,164	–0.9503	0.0088	0.0124	0.9999	0.1098	–1.4882	45.3789
Denmark	249	1,230	1,479	–0.9976	0.0087	0.0154	0.9559	0.1037	–4.6065	54.6666
Finland	165	775	940	–0.9995	0.0127	0.0152	0.3309	0.0658	–5.5759	76.7922
France	1,519	6,009	7,528	–0.9920	0.0064	0.0138	0.9819	0.0966	–3.7171	46.3569
Germany	1,320	4,448	5,768	–0.9968	–0.0034	0.0076	0.9918	0.0894	–4.0841	43.4740
Greece	134	716	850	–0.9078	0.0274	0.0240	0.7055	0.0842	–0.3889	40.7506
Ireland	258	708	966	–0.9899	–0.0486	0.0189	0.7526	0.2201	–2.5592	9.7884
Italy	501	1,694	2,195	–0.9996	0.0073	0.0135	0.9999	0.0889	–4.2463	61.8068
Luxembourg	27	87	114	–0.2901	0.0236	0.0176	0.3872	0.0739	0.6656	11.8584
Netherlands	341	2,009	2,350	–0.9852	0.0119	0.0162	0.8244	0.0769	–5.9090	76.2133
Portugal	47	225	272	–0.5868	0.0174	0.0147	0.3533	0.0863	–1.6497	21.1799
Spain	184	1,177	1,361	–0.8078	0.0318	0.0210	0.9088	0.0797	0.2727	33.9162
Sweden	372	1,432	1,804	–0.9978	–0.0004	0.0172	0.9683	0.1332	–3.2069	27.2618
UK	6,579	18,769	25,348	–0.9999	–0.0221	0.0183	0.9847	0.1737	–3.2141	16.1894

Note: The final working sample is free from zero Sales and zero Costs; zero Earnings (Austria 2, France 2, Spain 5, UK 8, US 26) are included with Profits for this tabulation. The measures of skew and kurtosis are non-standardised – given the r^{th} central moment $m_r = 1/n\Sigma(x_i - x_{mean})^r$ for $i=1,2,\ldots,n$, then skewness is defined as $m_3 m_2^{-3/2}$ and kurtosis as as $m_4 m_2^{-2}$.

Figure 3
Density estimation for scaled earnings, by jurisdiction

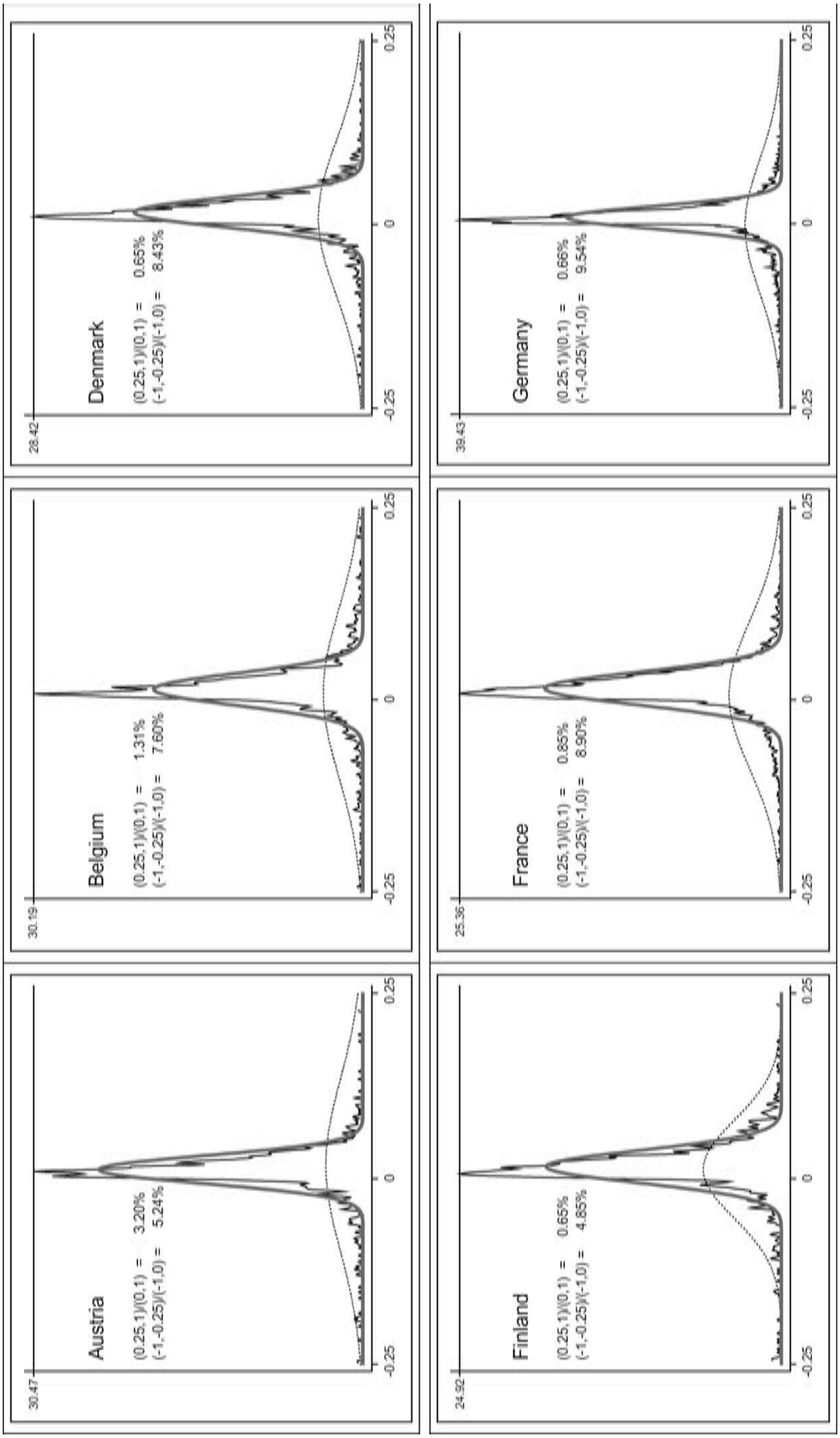


Figure 3
Density estimation for scaled earnings, by jurisdiction (*continued*)

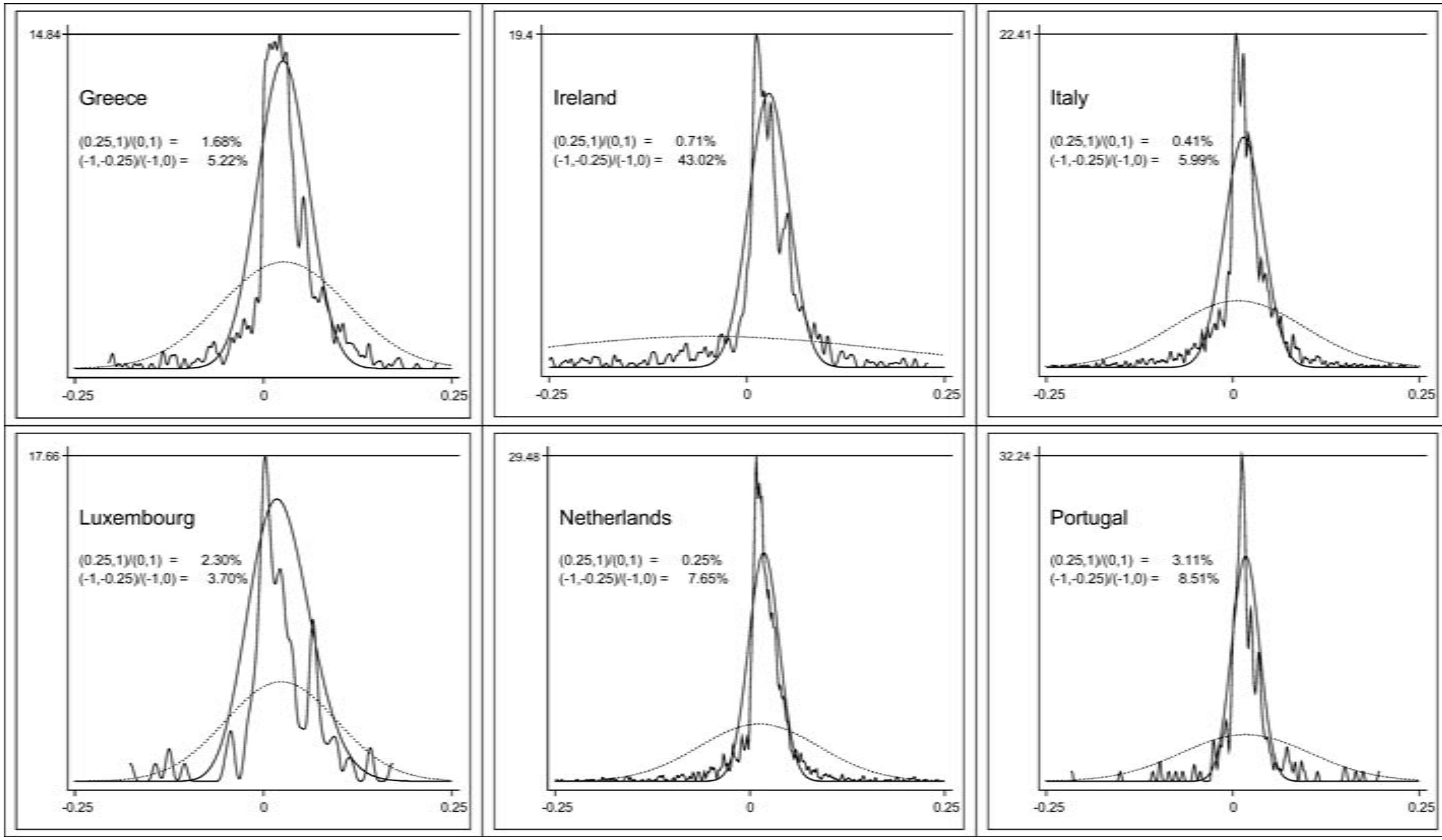
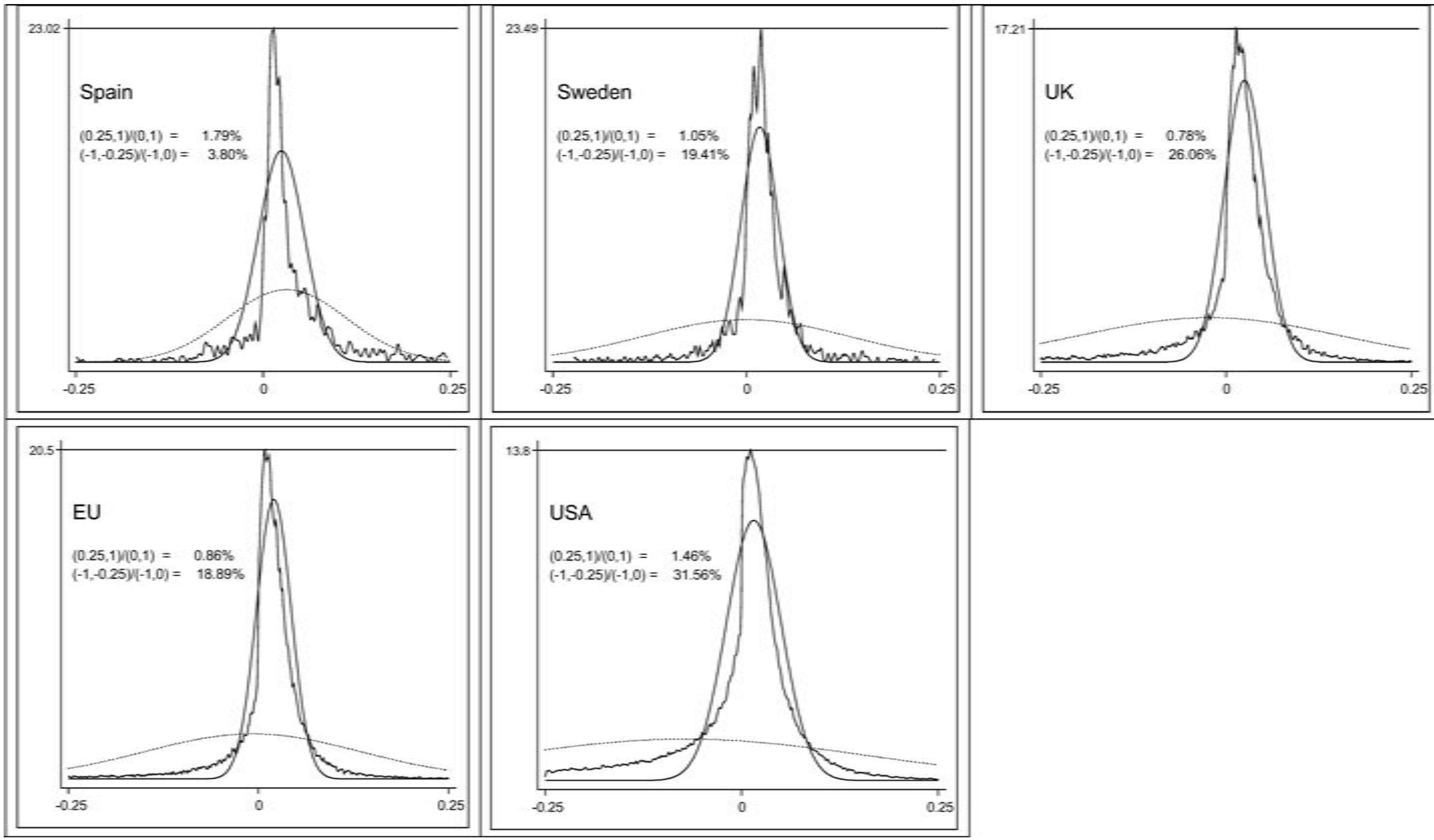


Figure 3
Density estimation for scaled earnings, by jurisdiction (*continued*)



Note: The black line is the kernel density estimator of E' defined by the Parzen (1962) kernel function with the Silverman (1986) bandwidth, the dotted curve is a normal density $N(\mu,\sigma)$ fitted on the respective sampling moments (see Table 1), and the grey curve is the fitted bounded distribution. The samples are free from observations with zero Sales and zero Costs. The y-axis labels indicate the maximum probability density. (0,0.25)/(0,1) and (–0.25,0)/(–1,0) indicate the percentage of non-visible profit and non-visible loss observations to the total number of profits and losses, respectively.

there is remarkable consistency in earnings behaviour across the EU.

This is evident in Figure 3, which juxtaposes the three density estimators of Scaled Earnings: the nonparametric Parzen kernel estimator, the bounded distribution (solid smoothed line) and the normal (dotted smoothed line). Each of these is fitted to the EU and US samples, and separately for each of the 15 member states of the pre-enlargement EU. While estimation is applied to the entire space of E', i.e. [*min*,*max*] for the kernel estimator and the normal and $[\xi,\xi+\lambda]$ for the bounded distribution, the graphs reproduced here focus on the interval [–0.25,0.25] in order to assist visual inspection about point zero. The consistently asymmetric shape of Scaled Earnings across the different jurisdictions is apparent in these plots, suggesting that Ijiri's (1965) characterisation of the zero point in accounting earnings as the modulator of asymmetry remains valid.

This asymmetric tendency in earnings can be further understood by looking at the asymmetric tail concentration between losses and profits, as reported by Figure 3. That is, in each separate plot, the concentration of out-of-range profit observations, for which $E' > 0.25$ is given as a percentage of the total number of profit observations, and the same approach is taken in order to calculate the concentration of out-of-range losses for $E' < -0.25$. It can be seen that, for both the EU and the US, the proportion of loss observations for which E' is less than -0.25 (EU 18.9%, US 31.6%) greatly exceeds the proportion of profit observations for which E' is greater than 0.25 (EU 0.9%, US 1.5%). This pattern is repeated in all of the member states confirming that, in all jurisdictions, loss observations tend to occur throughout their entire permissible region whereas profits do not, leading us to conclude that the asymmetry around its defining point of zero is a universal property of Net Income.[19]

Figure 3 also shows that the unbounded normal distribution fails systematically to fit the observed frequencies, which is not surprising given the inability of the Gaussian function to take into consideration the higher order moments that are required to define the general shape of earnings. The lack of fit in the case of the US provides a good illustration of the way in which the heavier tail density for losses shifts the normal's estimated point location downwards and away from the empirical mode. In all samples, there is considerable overfitting in the shoulders of the distribution, which arises because the normal cannot model the high peakedness that is characteristic of earnings.[20] In contrast, the bounded distribution is able to accommodate much of the shape of Scaled Earnings, in line with the description provided by the kernel density estimator. The Lagrange Multiplier test reported in Table 2 verifies that the deviance between the fit of Equation (8) and the observed data is significantly less than in the case of the normal fit (the test is described in the Appendix).

A parametric description of the scaled earnings distribution

Table 2 gives the recovered lower bound ξ, upper bound λ, scale $\xi+\lambda$, and shape parameters γ and δ, for the EU and US samples and additionally by member state of the EU. Table 2 also provides the fitted point estimates for the mean $\mu_1'(E')$, the median E_0' and the mode E_M', as well as the standardised median value $(E_0' - \xi)/\lambda = (1 + \exp(\gamma/\delta))^{-1}$ and the proportional distance of the median and mode from the mean $d = \left(E'_0 - \mu'_1(E')\right)/\left(E'_M - \mu'_1(E')\right)$.

For the bounded space of E', the optimisation process yields exact fits to the theoretical lower bound of reporting zero Sales (E'=–1) and to the theoretical upper bound of reporting zero Costs (E'=1), both for the single European market and the US. For the smaller subsamples by member state of the EU, the parametric fits take advantage of as much of the permissible range of variation as possible, as the iterative routine simultaneously solves for all parameters. These numerical results yield bounds of E' that are theoretically sound, the

[19] We have also applied Equation (5) to operating income and pre-tax income. The results, which are not tabulated in the paper, show that the concentration around zero and the distribution across the bounded range of variation are each sensitive to the level of income that is used. For scaled operating income, the distribution extends more smoothly to both tails and also passes through zero less abruptly. In the case of scaled pre-tax income, the inclusion of non-operating and transitory items, and prior period value adjustments, tends to increase the probability of a scaled profit and induces a more uniform allocation of scaled losses over their range. Finally, it is only with scaled net income, which takes account of all tax-related charges, plus minority interests and preferred dividend payments, that zero emerges clearly as the defining point of asymmetry in E', with the tax-effect mainly pulling profits away from the right-hand limit towards zero. Additional sensitivity analysis was also performed on the aggregation of all revenue items in addition to sales, and the results suggest that the distributional shape of (Revenues – Expenditures) / (Revenues + Expenditures) is similar to that of the more narrowly defined (Sales – Costs) / (Sales + Costs) reported in the paper.

[20] An additional test (Shapiro-Wilk-Royston), which is not tabulated in the paper, is overwhelmingly in favour of non-normality. Results by industry show that the industry factor to be much less influential than jurisdiction in shaping localised variability, with the exception of some observable differences between the cyclical and non-cyclical sectors of the economy.

Table 2
Recovered parameters for scaled earnings

	Parameters					*Estimates of central tendency*					*Model fitting*	
Pooled for 1985–2004	*Lower bound* ξ	*Scale* λ	*Upper bound* $\xi+\lambda$	*Shape* γ	δ	*Mean* $\mu_1'(E')$	*Median* E_0'	*Mode* E_M'	*Standardised median* $(E_0'-\xi)/\lambda$	*Proportional distance* $d\%$	*Lagrange Multiplier test*	*Optimal objective function*
EU	–1	2	1	–1.017	21.93	0.02317	0.02318	0.02321	0.5116	33.29	45,058.5	W-NM
US	–1	2	1	–0.256	15.19	0.00843	0.00844	0.00846	0.5042	33.24	39,883.2	W-NM
EU, by member state												
Austria	–0.8359	1.7480	0.9121	1.382	23.37	0.01228	0.01227	0.01224	0.4852	33.29	890.83	O-NM
Belgium	–0.9847	1.9847	1	–0.277	23.68	0.01344	0.01344	0.01345	0.5029	33.29	1,086.78	O-LM
Denmark	–0.9979	1.9979	1	–0.765	24.69	0.01652	0.01653	0.01654	0.5077	33.30	1,427.71	O-LM
Finland	–1	2	1	–0.724	22.84	0.01584	0.01585	0.01586	0.5079	33.29	849.87	O-LM
France	–0.9922	1.9922	1	–0.509	23.18	0.01483	0.01484	0.01485	0.5055	33.29	7,056.75	O-LM
Germany	–0.9969	1.9969	1	–0.453	32.86	0.00843	0.00843	0.00843	0.5034	33.31	5,236.61	O-LM
Greece	–0.9079	1.7331	0.8251	–1.989	12.78	0.02584	0.02595	0.02615	0.5388	33.20	748.55	O-LM
Ireland	–1	1.9928	0.9928	–1.248	19.85	0.02769	0.02771	0.02775	0.5157	33.28	524.92	W-NM
Italy	–1	2	1	–0.564	19.41	0.01452	0.01453	0.01455	0.5073	33.28	2,002.07	O-LM
Luxembourg	–0.2902	1.2902	1	6.666	5.832	0.02351	0.02173	0.01812	0.2418	33.06	87.11	D-NM
Netherlands	–1	1.9791	0.9791	–1.463	25.37	0.01806	0.01807	0.01810	0.5144	33.30	2,209.19	O-LM
Portugal	–1	1.9377	0.9377	–2.617	26.17	0.01724	0.01725	0.01729	0.5249	33.30	235.03	W-LM
Spain	–0.8545	1.7860	0.9315	0.457	14.58	0.02454	0.02452	0.02449	0.4922	33.23	966.03	O-LM
Sweden	–0.9981	1.9981	1	–0.672	20.70	0.01715	0.01716	0.01718	0.5081	33.28	1,631.73	O-LM
UK	–1	1.9980	0.9980	–0.945	18.18	0.02495	0.02497	0.02501	0.5130	33.27	18,946.30	W-NM

Notes: The parameter estimates indicate whether or not the theoretical limits of E' $\xi=-1$ and/or $\xi+\lambda=1$ were fitted; γ and δ are the fitted shape parameters. The estimates of central tendency are the expected mean (first moment) $\mu_1'(E')$, the median E_0', the uni-mode E_M', the standardised median $(E_0'-\xi)/\lambda=(1+\exp(\gamma/\delta))^{-1}$, and the proportional distance between the point location estimates $(E_0'-\mu_1'(E'))/(E_M'-\mu_1'(E'))$ which is positive when the median falls between the mean and the mode. The Lagrange Multiplier test, which examines the null that the fit of the bounded model to the observed data is the same as the fit of the normal to the same data, asymptotically converges to the χ_2^2 distribution with 2 degrees of freedom (all p-values are found to be less than 0.0001). The final column indicates the optimal objective function for the particular sample (ordinary, weighted or diagonally weighted least squares), and whether the Levenberg-Marquardt (LM) algorithm was successful in finding the optimal solution or gave way to the Nelder-Mead (NM) algorithm for the final convergence. Most samples are fitted by minimising either the O or the W objective functions, except Luxembourg which is distributed within a much narrower range of variation with significant weight in both tails and requires diagonally weighted least squares.

solution being independent from data points with exact contact at the limits (as noted earlier, we exclude observations with $E'=-1$ and $E'=1$).

We use the standardised median value $(E_0' - \xi)/\lambda = (1 + \exp(\gamma/\delta))^{-1}$ to compare the entire set of parametric fits $\{\xi, \lambda, \gamma, \delta\}$ across samples. The standardised median represents a sigmoid (or standard logistic) function of γ and δ, with range [0,1] and a cut-off point of $(1 + \exp(\gamma/\delta))^{-1}=0.5$ when $\gamma/\delta=0$. That is, for $\gamma=0$ the fitted distribution is symmetric. In the case of positive skewness, $(1 + \exp(\gamma/\delta))^{-1}\to 0$ as $\gamma/\delta\to\infty$, while for negative skewness, $(1 + \exp(\gamma/\delta))^{-1}\to 1$ as $\gamma/\delta\to-\infty$. The standardised median estimates are remarkably similar, with the exception of three relatively small member states which have positive γ (Austria 0.4852, Luxembourg 0.2418, and Spain 0.4922). For these jurisdictions, positive γ seems to be appropriate and simply reflects the range of their observed values which, in contrast to the other subsamples, range further over profits than losses (Austria: min=–0.7815, max=0.9015; Luxembourg: min=–0.2901, max=0.3872; Spain: min =−0.8078, max=0.9088).

Table 3
Filliben's percentile correlation coefficient

Pooled for 1985–2004	*Johnson transformation*	*Normal fit*
EU	0.950	0.507
US	0.963	0.546
Austria	0.969	0.529
Belgium	0.962	0.572
Denmark	0.967	0.584
Finland	0.957	0.723
France	0.963	0.612
Germany	0.956	0.547
Greece	0.949	0.759
Ireland	0.968	0.453
Italy	0.950	0.682
Luxembourg	0.914	0.801
Netherlands	0.959	0.616
Portugal	0.906	0.715
Spain	0.963	0.550
Sweden	0.970	0.531
UK	0.969	0.529

Notes: The critical values are interpolated from Vogel (1986) as follows: *normal distribution*: 0.978 (0.1%), 0.979 (0.5%), 0.981 (1%) 0.987 (5%); *extreme value (Type I) distributions*: 0.949 (0.1%), 0.952 (0.5%), 0.960(1%), 0.978 (5%).

The fitted mean, median and mode reported in Table 2 are more robust than their nonparametric counterparts reported in Table 1, the sample mean and the sample median. Appropriately, these fitted estimates are always positive, consistent with the sign of expected earnings in a viable economy, and the median always falls between the mean and the mode and thus reflects the continuous unimodal density that has been proposed. In contrast, the sensitivity of the arithmetic average to extreme values in the sample can be seen to lead to negative sample means in both the EU (–0.0077) and US (–0.0712).[21] This is a severe shortcoming of relying on nonparametric estimates – given the considerable number of firms and years involved, it would be implausible that the most likely expected value of earnings, in either the EU or the US, would be a loss.

The fitted estimates of central tendency also explain how the different tail weights give rise to high density just above zero and negative skew. This is a consistent result that becomes particularly evident when we look at *d*, the distance of the median from the mean $\left(E_0' - \mu_1'(E')\right)$ expressed as a percentage of the distance of the mode from the mean $\left(E_M' - \mu_1'(E')\right)$. This must be positive when the median of a distribution falls between the mean and the mode. Two inferences may be drawn from the analysis of the proportional distance. First, $\left(E_0' - \mu_1'(E')\right)$ and $\left(E_M' - \mu_1'(E')\right)$ are always very small (i.e. a difference is only observed at the fourth decimal place), which reflects the high concentration just above zero as well as the expectation that the most likely value is indeed a small scaled profit. Second, there is a remarkable similarity in proportional distance across subsamples, consistently estimated close to 33%. In other words, the distance between the mode and the median is twice as large as the distance between the median and the mean in all cases, even for subsamples that are positively skewed, with expected concentration just above zero following a consistent pattern throughout.

The model fitting diagnostics reported in the final columns of Table 2 provide compelling support for the Johnson transformation. The Lagrange Multiplier test strongly favours the fit

[21] The fitted distributions comply with the Dharmadhikary and Joag-dev (1988) conditions, under which the median must fall between the mode and the mean for a finite continuous unimodal distribution (see also: Basu and DasGupta, 1997; Bickel, 2003). These conditions ensure that mean≤median≤mode under negative skewness, and that mean≥median≥mode under positive skewness (as with the three EU sub-samples of Austria, Luxembourg and Spain). This relationship does not always hold for the sample mean and median, simply because these are nonparametric results that are highly sensitive to any form of abnormality (e.g. Greece mean=0.0274 $>$ median=0.0240 but skewness=–0.3889; Netherlands mean=0.0174 $>$ median=0.0147 but skewness=–1.6497).

of the bounded model to the observed data over the fit of the normal to the same data (all p-values are found to be less than 0.0001). The test, which is specified in the Appendix (Equation A5), converged on an optimal solution using standard techniques in all cases except for the Luxembourg subsample, which is by far the smallest (N=114) and which is distributed empirically within a much narrower range of variation. Finally, indicative goodness-of-fit statistics for the Johnson transformation and the normal are documented in Table 3, using an extension of Filliben's quantile-quantile correlation test. Standard goodness-of-fit measures are designed for relatively small randomised samples, whereas the commercial data sets that are common in accounting research provide population coverage, and therefore tend to be very large. They require a different approach, especially in order to compare across jurisdictions. Hence the use here of the Filliben correlation test. If the sample is distributed as hypothesised, we expect the relation between the ordered cumulative distribution function (CDF) to be linear with respect to the theoretical CDF, and similarly for the i^{th} order statistics. In this case, the product moment correlation between the percentiles of the empirical cumulative frequencies and the theoretical cumulative frequencies provides the appropriate statistic, which has the advantage of being applicable to distributions other than the normal (Vogel, 1986; Heo et al., 2008). The results of this indicative test show high association with the best-fitting Johnson transformation (average 0.95, minimum 0.91) and much lower association with the fitted normal (average 0.61, minimum 0.45). Indeed, in all cases except the two smallest samples, the hypothesis that the observed data fits the Johnson transformation cannot be strongly rejected, whereas for the normal distribution the hypothesis is rejected outright.

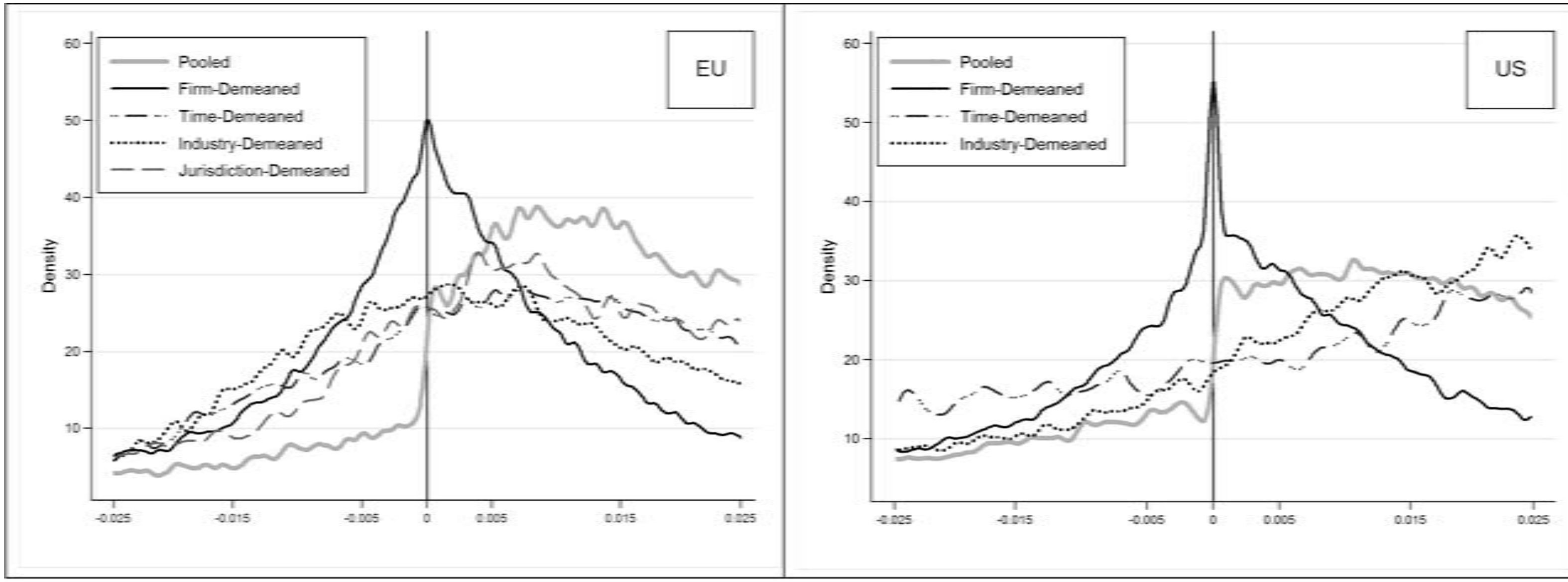
The goodness-of-fit tests indicate the appropriateness of the models that are fitted, and the parametric analysis set out above validates our claims for a consistently asymmetric shape of earnings that is chiefly described by negative skewness and high levels of concentration just above zero. On the whole, there is little difference in the shape of fitted densities across samples. More specifically, it is shown how asymmetry in scaled earnings is primarily defined by a longer tail for losses and a shorter tail for profits, with zero acting as Ijiri predicted, modulating the downwards pressure not only on profits, which is evident in the high density just above zero, but also on losses resulting in lower concentration just below zero.

Is the asymmetry about zero a firm-specific effect? Finally, we provide evidence that suggests that the asymmetry in earnings may be a feature that is predominantly introduced through firm-specific heterogeneous effects. The income of firms in a competitive environment will be attributable to the characteristics of each entity but conditional in each case on the firm's relationship with the rest of the market. If markets are complete, with perfect information and homogeneity in the allocation of resources, the reported firm profit in this frictionless universe will be absent of incentives and other stimuli that create asymmetry. We anticipate therefore, that once we remove the fixed effect that defines firm-specificity, we will induce approximate symmetry about zero. The standard approach in panel methods is employed here, whereby the arithmetic averages for each panel (in this case a firm) are the fixed effects. Such effects are described in modern microeconometric analysis as either unobserved or unobservable, and they characterise the between-panel heterogeneity in the pooled sample (e.g. Cameron and Trivedi, 2005).

To examine this proposition, consider a more comprehensive description of Scaled Earnings E' with fixed attributes for each firm i, for each of the sectors s in which the sampled firms operate, for each of the jurisdictions j in which they are domiciled, and for each year t. The fixed attributes are thus the firm mean $\bar{E}'_i$ for i=1,2,…,n, the sector mean $\bar{E}'_s$ where s denotes a two-digit SIC class, the jurisdiction mean $\bar{E}'_j$ where j denotes an EU member state, and the year-by-year mean $\bar{E}'_t$ for each t=1985,1986,…,2004. By subtracting the arithmetic average of the earnings stream with respect to any one of these attributes, we effectively eliminate the expected heterogeneous effect that is associated with that particular trait. We then repeat the Parzen-kernel estimation procedure with the mean-adjusted data, in order to assess which, if any, of these characteristics may cause the distribution to deviate from symmetry.

Panel A of Figure 4 contrasts the kernel density estimators of the pooled firm-year sample of Scaled Earnings E' with the densities of observations that are mean-adjusted by firm $\left(E' - \bar{E}'_i\right)$, by sector $\left(E' - \bar{E}'_s\right)$, by year $\left(E' - \bar{E}'_t\right)$ and, in the case of the EU only, by jurisdiction $\left(E' - \bar{E}'_j\right)$. As we are mainly interested in observing the distribution of earnings around zero, the graph focuses on the central five percentiles, i.e. over the range [–0.025, 0.025]. The effect is very noticeable in both the EU and the US, suggesting that heterogeneity across firms is the main cause of asymmetry in the pooled samples. By comparison with the pooled data,

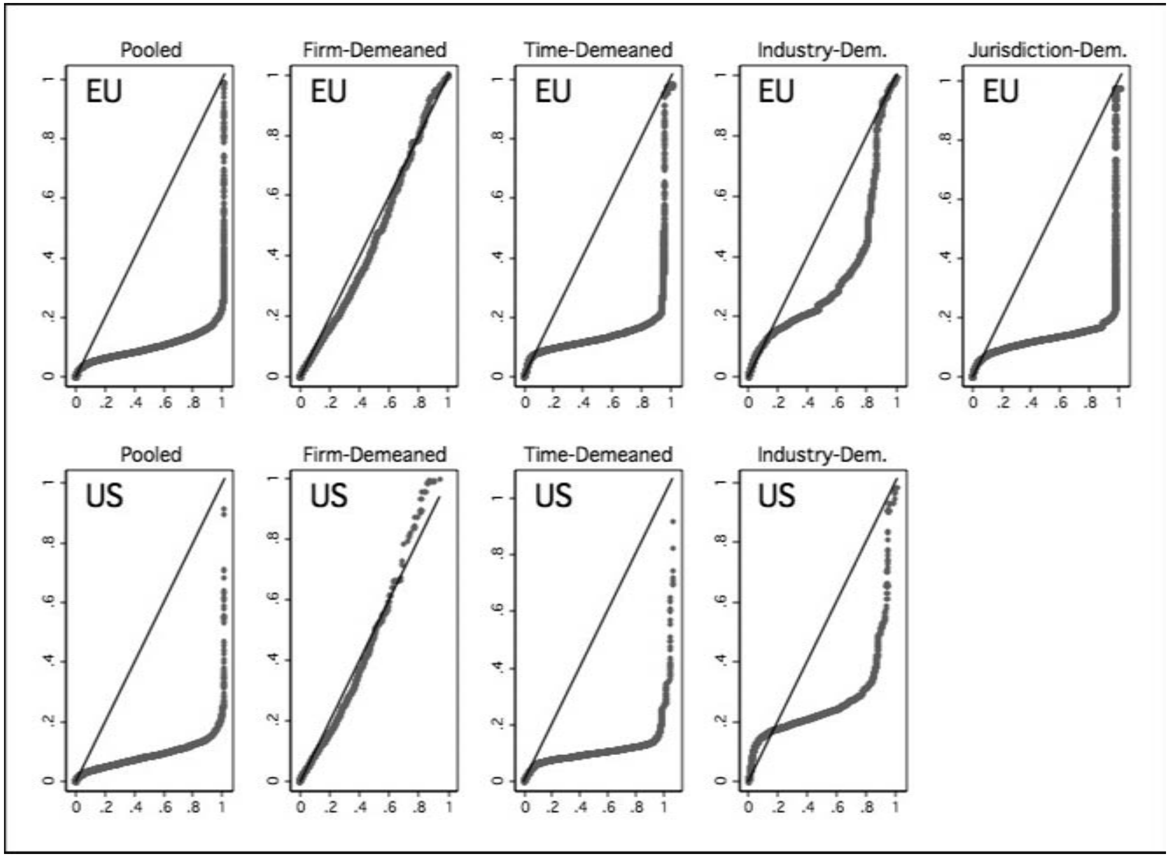
Figure 4
Panel A: Asymmetry in earnings is a firm-mean effect



	Unadjusted	Mean adjusted			
	Scaled Earnings E′	*by firm*	*by sector*	*by year*	*by jurisdiction*
EU *standard deviations*	0.1420	0.0851	0.1384	0.1398	0.1410
Levene test	Mean-adjusted vs. *E′*	*p*<0.0001	*p*<0.8595	*p*<0.5121	*p*<0.3979
US *standard deviations*	0.2299	0.1287	0.2134	0.2281	n/a
Levene test	Mean-adjusted vs. *E′*	*p*<0.0001	*p*<0.0180	*p*<0.8067	n/a

Note: The kernel density plots above focus on the central 5% [–0.025,0.025]. The estimates for the pooled unadjusted Scaled Earnings E' are plotted together with those for the mean-adjusted values by firm$\left(E' - \bar{E}'_i\right)$, by sector $\left(E' - \bar{E}'_s\right)$, by year $\left(E' - \bar{E}'_t\right)$ and, for the EU only, by jurisdiction$\left(E' - \bar{E}'_j\right)$. The data set comprises of 53,087 firm-years in the EU and 134,234 in the US. The table above gives the standard deviations for each of the unadjusted and mean-adjusted samples, together with the Levene Test for Homogeneity of Variance, which is a robust test for examining the null of equality in variances between non-normal frequencies, and follows an F distribution with (1, N–2) degrees of freedom.

Figure 4 (*continued*)
Panel B: Symmetry plots



Note: The symmetry plots above show the distance from the median of the observed distribution, i.e. the *y*-axis measures the distance above the median and the *x*-axis measures the distance below the median. The plot region has been restricted to the interval [–1,1] for both axes.

where the standard deviation is $\sigma = 0.1420$ in the EU and 0.2299 in the US, the mean-adjusted density at the firm level shows considerably less variability (σ = EU 0.0851, US 0.1287), which is not the case for the other mean adjustments where the reduction in variability does not appear to be material (by sector, σ = EU 0.1384, US 0.2134; by year, σ = EU 0.1398, US 0.2281; by jurisdiction in the EU, $\sigma = 0.1410$ – see Figure 4). We test for the equality of variances between the unadjusted frequencies and the mean-adjusted frequencies using the Levene (1960) Test for Homogeneity of Variance, which is a robust test for non-normal frequencies.[22] As shown by the tabulation accompanying Panel A of Figure 4, the results confirm that the variances are indeed significantly different between the unadjusted and the firm mean-adjusted frequencies, with p-values less than 0.01%. On removing the other fixed effects that are considered here, the level of dispersion tends to remain statistically unaltered with the exception of the sector effect for the US, implying that, in the US, there are stronger industry-specific effects than in the EU (at least at the level of two digit SIC codes).[23]

The symmetry plots of Scaled Earnings at the pooled and mean-adjusted levels in Panel B of Figure 4 show the distance from the median of the observed distribution. As can be seen, in both the US and the EU, the pooled distribution is highly asymmetric, as are the mean-adjusted densities with respect to time and industry (and jurisdiction in the case of the EU), but the mean-adjusted distribution at the firm level is extremely close to symmetry. In summary, the induced symmetry in earnings that is achieved by removing the heterogeneous fixed effect at the firm level substantiates our earlier assertion that earnings should not be viewed as a mixture of distributions between firms that make losses and firms that make profits. More importantly, it verifies the claim that earnings asymmetry is attributable to firm-specific factors, and again validates Ijiri's notion that zero acts both as a constraint and an objective for the management of the firm, modulating asymmetry between reported losses and profits.

Is it possible that this result arises just because the variance in firm averages is considerably greater than the variance in industry or year averages? This is clearly a key issue for further research. This study has demonstrated that symmetry can be approximated at the firm level, suggesting that the factors that give rise to asymmetry are indeed firm-specific and are not related to time, industry or jurisdiction.[24] It should be noted, however, that we have applied a time-invariant fixed-effects model to the firm-year panel, without extending the analysis to random effects or dynamic modelling. In our analysis, the remaining disturbance varies with both firms and years, and is treated as the IID random component. Further investigation calls for the application of more detailed panel methods of analysis. It is worth emphasising here, by way of conclusion, how little impact panel methods, even fixed effects models, have had on accounting research to date. The perceived need to delete the upper and lower percentiles of observations (on the grounds that they are possibly outliers) appears to be ingrained as the standard approach with cross-sectional methods, and this surely places a major constraint on the adoption of panel structures in accounting research – the ad hoc removal of observations from firm-specific time series is not consistent with robust panel estimation. The bounded scalar proposed in this paper goes some way towards resolving this issue, as such ad hoc approaches would then become redundant. Clearly, more work is needed to untangle the basic properties of accounting variables in firm-year panel datasets.[25]

7. Concluding remarks

This paper is motivated by prior research that has suggested a 'discontinuity' about zero in the

[22] The Levene test examines the null of equality in variances. For a variable X, with total sample size N, and sub-samples N_1 and N_2, the Levene test follows an F distribution with (1, N–2) degrees of freedom and is defined as

$$L = (N-2)\left(\sum_{j=1}^{2} N_j\left(\bar{W}_{.j} - \bar{W}_{..}\right)^2 / \sum_{j=1}^{2}\sum_{i=1}^{N_j} N_j\left(\bar{W}_{ij} - \bar{W}_{.j}\right)^2\right) \sim F^{\alpha}_{1,N-2}$$

where $\bar{W}_{ij} = \left|\bar{X}_{ij} - \bar{X}_{.j}\right|$, $\bar{X}_{.j}$ is the arithmetic mean, $\bar{W}_{.j}$ the group means of $\bar{W}_{ij}$, and $\bar{W}_{..}$ the global mean of $\bar{W}_{ij}$ (Brown and Forsythe, 1974).

[23] By removing the firm-median instead of the firm-mean, we find that the transformation to symmetry and the reduction in variance is even more powerful. This is not surprising, since the distribution of earnings is fundamentally non-normal with significant higher order moments. The heterogeneous firm-specific effects described here play a decisive role in empirical relationships involving accounting variables, as demonstrated in the assessment of fixed and random effects in earnings conservatism in Grambovas et al., (2006).

[24] Exploratory simulations appear to confirm that asymmetry is a firm-specific effect. For example, by randomising the real data to form random firm-panels, e.g. of 10 or 20 years, from which the mean is then deducted, it is evident that the mean-adjusted frequencies remain highly asymmetric. This is the case for random samples with similar sample properties to the empirical observations, as defined earlier in footnote 10, and suggests again that asymmetry is shaped predominantly by firm-specific fixed effects and not random effects.

[25] By way of introduction, we refer the reader to the panel analysis of earnings in Grambovas et al., (2006), which provides a detailed discussion of fixed and random effects models, and to the panel analysis of ratios of two scalars in McLeay and Stevenson (2008). Extending the discussion of fixed and random effects into the context of econometric modelling of accounting data opens up a wide agenda, including: the presence of joint unit roots in accounting variable series, the

distribution of accounting earnings, and recent suggestions that the earnings scalar that is selected may bias in favour of finding such discontinuities. In the introduction, we have demonstrated how previous research using a histogram estimator to define the weights of observed probabilities may be influenced by the way in which the observed data are aggregated, as asymmetry in Scaled Earnings observations around zero also tends to be accentuated by the choice of histogram origin and bin width.

To examine the shape of the distribution of scaled earnings, in this paper we use both nonparametric and parametric density estimators. The kernel density estimator provides a detailed and unbiased nonparametric description of localised variability, which is independent from discrete choices on groupings of data, and the Johnson bounded distribution provides an appropriate parametric model that is consistent with our scaling approach. We begin the analysis of the earnings distribution by proposing a simple bounded model, which scales earnings over a restricted range of variability. This model expresses earnings as the percentage return on the total magnitude of transactions that flow within a financial year, and has limits that are defined by the extreme reporting behaviour of zero Sales and zero Costs.[26] The proposed model avoids some of the statistical shortcomings usually arising with earnings scalars, such as infinite variances and extreme observations, and rotates around zero with deviations restricted to no more than one standard unit. Moreover, the scalar introduced in this paper, Sales+Costs, will correct for this bias because losses can be attributed not only to falling Sales but also to increasing Costs.

Across two major economic regions, the EU and the US, and by member state jurisdictions within the EU, both the nonparametric and parametric results confirm a consistent pattern of non-normality in the form of asymmetric expected variance between loss and profit observations, with great concentration just above zero. It is demonstrated in this paper that the inherently non-normal pattern could be mainly the result of firm-specific factors.

In the earnings management literature, expectations based on assumptions of symmetry have led to confusing statements along the lines that ‘there are more small profit observations than expected and fewer small loss observations than expected’, leading to the interpretation of such findings as prima facie evidence of earnings management. However, in this paper, we argue against the null hypothesis of a distribution that is smooth in the region around zero, and instead expect an inherent asymmetry in profits and losses.

setting of initial conditions for censored panels where exponential growth is a characteristic of firm data, the placement of autocovariance restrictions on variables resulting from double-entry book-keeping, and the potential for non-linear systems, error-correction features and structural equations that may jointly represent a variable of interest and its scalar when these are strictly linked by accounting identities.

[26] The range of scaled earnings is [–1,1]. The extreme cases are uncommon, but not ‘abnormal’, as there are real factors underlying non-trading or nonproduction in a given period. The causes of such extreme reporting behaviour are not analysed in this paper. Nevertheless, the data used in this study are publicly available, and therefore this phenomenon is readily observable: a listing with the names of all of the sampled firms in the EU that have reported zero Costs and/or zero Sales in the study period may be obtained from the authors. Further work on understanding the nature of zeros in accounting, and their treatment in empirical analysis, would make a useful contribution to the research literature.

Appendix
Recovering the fitted parameters for scaled earnings

For a continuous variable E' that is randomly drawn from an IID sample, then for known ξ=–1 and λ=2 and the logit transformation $g_{it} = \ln\left(\left(E'_{it} - \xi\right)/\ \left(\xi + \lambda - E'_{it}\right)\right)$, the maximum likelihood estimates for the shape parameters of Equation (8) are $\tilde{\delta} = 1/s_g$ and $\tilde{\gamma} = -\tilde{\delta}\bar{g}$, where $\bar{g}$ is the arithmetic sample mean and s_g the sample standard deviation of g_{it} (Johnson, 1949). However, it is known that maximum likelihood estimation systematically fails when there is high concentration around the mean (Kottegoda, 1987; Siekierski, 1992; Zhou and McTague, 1996). Since we expect such high levels of kurtosis in the distribution of Scaled Earnings E', for the purposes of this paper we turn to the more flexible least squares methodology developed by Swain et al. (1988) to address this particular concern.

The quadratic problem of least squares seeks to minimise the sum of squared ordinate differences between the ranked parametric approximation of the fitted bounded distribution to the standard normal $\gamma + \delta g_{(it)} \sim z(0, 1)$ and the nonparametric expected standard normal scores $i/(n + 1)$, as follows:

$$\sum_{it=1}^{N} \varepsilon_{it}^2 = \sum_{it=1}^{N} \left(\Theta\left\{\gamma + \delta \ln\left(\frac{E'_{(it)} - \xi}{\xi + \lambda - E'_{(it)}}\right)\right\} - \frac{it}{N+1}\right)^2, \tag{A1}$$

where $\Theta\{.\}$ indicates the translation to z, and $E'_{(it)}$ the ascendingly ordered sample of $\xi < \{E'_{(1)} \leq E'_{(2)} \leq \ldots \leq E'_{(N)}\} < \xi + \lambda$. By assuming independent and homoscedastic errors we minimise the sum of equally weighted $\sum \varepsilon_{it}^2$ by searching for the best possible OLS fit of $\{\xi, \lambda, \gamma, \delta\}$, under the following conditions:

$$\underset{\xi,\lambda,\gamma,\delta}{\overset{OLS}{\min}} \left\{ \sum_{it=1}^{N} \varepsilon_{it}^2 \middle| \begin{array}{c} -1 \leq \xi < E'_{(1)} \\ (E'_{(N)} - \xi) < \lambda \leq (1 - \xi) \\ \text{sign}\{\gamma\} = \text{sign}\{b_1\} \\ \delta > 0 \end{array} \right\} \tag{A2}$$

Equation (A2) gives the objective function with bounds for ξ, λ and δ, but allows γ to take any value or be equal to zero as long as it bears the same sign as the skewness coefficient b_1. The sign constraint on γ is a useful condition that directs optimisation so that it escapes local optima that are implied by initial values. With regard to the bounds, if ξ=–1 and λ=2 are representative extrema of the expected sample space, then only the conditions for γ and δ need to be satisfied. Yet, it has been shown that by allowing estimation of narrower limits within the bounded range it is possible to significantly improve the quality of the fit (Tsionas, 2001).

Alternatively to Equation (A2), Swain et al. (1988) suggest that when estimating a non-linear model, such as the one considered here, it may prove useful to relax the OLS assumptions and instead minimise the following weighted least squares (WLS):

$$\underset{\xi,\lambda,\gamma,\delta}{\overset{WLS}{\min}} \left\{ 2(N+1)(N+2)\left(\sum_{it=1}^{N} \varepsilon_{it}^2 - \sum_{it=2}^{N} \varepsilon_{it}\varepsilon_{it-1}\right) \middle| \mathit{conditions} \right\}, \tag{A3}$$

subject to the same *conditions* given for Equation (A2). In addition, Swain et al. (1988) further demonstrate how the inverse variances of the errors

$$Var(\varepsilon_{it})^{-1} = (N+2)/[it(N+1)^{-1}(1 - it(N+1)^{-1})]$$

generate a special type of objective function known as the diagonally weighted least squares (DWLS), as follows:

$$\underset{\xi,\lambda,\gamma,\delta}{\overset{DWLS}{\min}} \left\{ \sum_{it=1}^{N} Var(\varepsilon_{it})^{-1} \varepsilon_{it}^2 \middle| \mathit{conditions} \right\}, \tag{A4}$$

Appendix
Recovering the fitted parameters for scaled earnings (*continued*)

again, subject to the same *conditions* listed above. DWLS is the minimum variance linear unbiased estimator. In effect, it assigns larger weights to the tails of the ordered frequency in relation to the middle of the range, and therefore is expected to perform better in situations that require a good fit in the tails or the shoulders of the distribution.

We solve the optimisation routines of Equations (A2), (A3) and (A4) for each of the following four settings: (*i*) $\xi = -1$ and $\xi + \lambda = 1$, (*ii*) $\xi = -1$ and $\xi + \lambda < 1$, (*iii*) $\xi > -1$ and $\xi + \lambda = 1$, and (*iv*) $\xi > -1$ and $\xi + \lambda < 1$. The optimal solution conditional on the objective function and setting is the one with the minimum converged sum of squares, and the smallest L_∞-norm of the difference between the observed and the fitted density.

To solve these iterative problems we modify the program given by Swain et al. (1988) for fitting Johnson distributions (see also DeBrota et al., 1988). These modifications include setting starting values for the MLE shape parameters $\tilde{\gamma}$ and $\tilde{\delta}$ and the theoretical bounds of the distribution $\xi = -1$ and $\xi + \lambda = 1$, placing the tolerance of parameters at the fourth decimal place, and fitting large samples. We make use of both available iterative algorithms that were originally chosen for their design to address the specific problem of least squares. Optimisation starts by applying the Levenberg-Marquardt algorithm (LMA), which can be considered as a hybrid of the steepest gradient descent algorithm when it searches away from the global minimum, and the Gauss-Newton algorithm (GMA) when it approaches the global minimum. However, although the LMA is more robust than the GMA, it can still be very slow in converging conditional on the size of the problem, starting values, tolerance limits and shape of the terrain. For this reason, we allow for only a maximum number of 150 iterations and, if the LMA fails to converge, the optimality search switches to the Nelder-Mead algorithm (NMA) of the simplex class of methods for completion. The NMA is relatively robust and numerically uncomplicated since it does not require the evaluation nor the existence of derivatives, and therefore can converge more easily. Both algorithms have a downhill orientation which is appropriate for the minimisation problem of the sum of squares, i.e. a second order polynomial with zero ogees. Yet, since these algorithms may cause numerical instability we restart each procedure by using as initial values the previously converged set of parameters. We repeat this process until there is no change at the fourth decimal place.

It should be noted that, although the main aim for this paper has been to strike a balance in scaling over the peak and the shoulders, it is possible nevertheless to choose the density region which requires particular attention. For example, if a precise description of the peak is required, an objective function may be employed that minimises the distance between the theoretical and the observed global maximum (e.g. an L_∞-norm function), whereas a good fit in the shoulders may be achieved by solving a system of percentile equations that describe the particular region (e.g. Mage, 1980; Flynn, 2006). Further improvement may be achieved through enhanced optimisation routines, including algorithms for unsmooth functions with multiple local optima and unsuccessful starting values (see Cormen et al., 2001; Michalewicz and Fogel, 2002).

Comparing the fit to the normal

The fit of Equation (A2) can be tested against the alternative of a normal fit by restricting Equation (7) so that $\gamma + \delta E' = z \sim N(0,1)$. Accordingly, the errors are computed as follows:

$$\sum_{it=1}^{N} \varepsilon_{it}^2 = \sum_{it=1}^{N} \left(\Theta\left\{ \gamma + \delta E'_{(it)} \right\} - \frac{i}{N+1} \right)^2 \tag{A5}$$

We employ a Lagrange Multiplier test $\mathrm{LM} = n(SSE_R - SSE_U)/(SSE_R)$ to examine the restricted fit to the normal, where SSE_U and SSE_R are the sum of squared errors for the unrestricted Equation (A1) and the restricted Equation (A5), respectively. This LM test follows a χ^2 distribution with 2 degrees of freedom that represent the number of parameter restrictions placed in Equation (7), i.e. $\xi = 0$ and $\lambda = 1$.

Appendix
Recovering the fitted parameters for scaled earnings (*continued*)

Parametric point estimates of central tendency

Expected mean: Although the non-central moments are cumbersome to compute, Johnson (1949) shows that the first moment may be expressed as a ratio of infinite series that are independent of integral calculations, as follows:

$$\mu_1' = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\gamma^2} \left(\frac{\frac{1}{2\delta}+\frac{1}{\delta}\left(\sum_{n=1}^{\infty} e^{-n^2/2\delta^2} \cosh\left(\frac{n(1-2\gamma\delta)}{2\delta^2}\right) \operatorname{sech}\left(\frac{n}{2\delta^2}\right)\right)}{1+2\sum_{n=1}^{\infty} e^{-2n^2\pi^2\delta^2}\cos(2n\pi\gamma\delta)} - \frac{2\pi\delta\left(\sum_{n=1}^{\infty} e^{-\frac{1}{2}(2n-1)^2\pi^2\delta^2} \sin((2n-1)\pi\gamma\delta)\operatorname{cosech}\left((2n-1)\pi^2\delta^2\right)\right)}{1+2\sum_{n=1}^{\infty} e^{-2n^2\pi^2\delta^2}\cos(2n\pi\gamma\delta)} \right) \tag{A6}$$

Given the fitted $\{\xi, \lambda, \gamma, \delta\}$ and the relatively large values of δ, we solve Equation (A6) for that n at which $e^{-n^2/2\delta^2} = e^{-\frac{1}{2}(2n-1)^2\pi^2\delta^2} = e^{-2n^2\pi^2\delta^2} = 0$ at double precision.

Expected median: As the bounded distributional function for Scaled Earnings E' can be written as a translation to the standard normal $z = \gamma + \delta \ln\left(\left(E' - \xi\right)/\left(\xi + \lambda - E'\right)\right)$, then by setting the median of z to zero, we derive the fitted median:

$$E_0' = \xi + \lambda(1 + \exp(\gamma/\delta))^{-1} \tag{A7}$$

Accordingly, we may re-write Equation (A7) in the form of a standardised median value $(E_0' - \xi)/\lambda = (1 + \exp(\gamma/\delta))^{-1}$.

Expected mode: The Johnson bounded system accommodates both unimodal and bimodal curves. By differentiating Equation (8) and equating to zero we obtain:

$$2\left(\frac{E' - \xi}{\lambda}\right) - 1 - \gamma\delta = \delta \ln\left(\frac{E' - \xi}{\xi + \lambda - E'}\right) \tag{A8}$$

For $\delta \geq 1/\sqrt{2}$, there is a single solution to Equation (A8), with a uni-mode E_M' which can be recovered using a non-linear root-finding algorithm (Kotz and Van Dorp, 2004).

References

Ashton, D., Dunmore, P. and Tippett, M. (2004). ‘Double-entry bookkeeping and the distributional properties of a firm’s financial ratios’. *Journal of Business, Finance & Accounting*, 31(5/6): 583–606.

Ataullah, A., Rhys, H. and Tippett, M. (2009). ‘Non-linear equity valuation’. *Accounting and Business Research*, 39 (1): 57–73.

Balanda, K. and MacGillivray, H. (1988). ‘Kurtosis: a critical review’. *The American Statistician*, 42(2): 111–119.

Ballantine, J., Forker, J. and Greenwood, M. (2007). ‘Earnings management in English NHS hospital trusts’. *Financial Accountability and Management*, 23(4): 421–440.

Basu, S. and DasGupta, A. (1997). ‘The mean, median and mode of unimodal distributions: a characterization’. *Theory of Probability and Its Applications*, 41(2): 210–223.

Beaver, W., McNichols, M. and Nelson, K. (2007). ‘An alternative interpretation of the discontinuity in earnings distribution’. *Review of Accounting Studies,* 12(4): 525–556.

Bickel, D. (2003). ‘Robust and efficient estimation of the mode of continuous data: the mode as a viable measure of central tendency’. *Journal of Statistical Computation and Simulation*, 73(12): 899–912.

Brown, L. and Caylor, M. (2005). ‘A temporal analysis of quarterly earnings thresholds: propensities and valuation consequences’. *Accounting Review*, 80(2): 423–440.

Brown, M. and Forsythe, A. (1974). ‘Robust tests for the equality of variances’. *Journal of the American Statistical Association*, 69(346): 364–367.

Burgstahler, D. and Dichev, I. (1997). ‘Earnings management to avoid earnings decreases and losses’. *Journal of Accounting and Economics*, 24(1): 99–126.

Burgstahler, D. and Eames, M. (2003). ‘Earnings management to avoid losses and small decreases: are analysts fooled?’. *Contemporary Accounting Research*, 20(2): 253–294.

Cameron, A. and Trivedi, P. (2005). *Miroeconometrics: Methods and Applications*. New York, NY: Cambridge University Press.

Cobb, L., Koppstein, P. and Chen, N. (1983). ‘Estimation and moment recursion relations for multimodal distributions of the exponential family’. *Journal of the American Statistical Association*, 38(381): 124–130.

Cooke, T. and Tippett, M. (2000). ‘Double entry bookkeeping, structural dynamics and the value of the firm'. *British Accounting Review*, 32(3): 261–288.

Cormen, T., Leiserson, C., Rivest, R. and Stein, C. (2001). *Introduction to Algorithms*, MIT Press.

Daniel, N., Denis, D. and Naveen, L. (2008). 'Do firms manage earnings to meet dividend thresholds?'. *Journal of Accounting and Economics*, 45(1): 2–26.

Daske, H., Gebhardt, G. and McLeay, S. (2006). 'The distribution of earnings relative to targets in the European Union'. *Accounting and Business Research*, 36(3), 137–168.

DeBrota, D., Roberts, S., Swain, J., Dittus, R., Wilson, J. and Venkatraman, S. (1988). 'Input modelling with the Johnson system of distributions'. *Proceedings of the 20th Conference on Winter Simulation*, San Diego, CA, December 12–14: 165–179.

Dechow, P., Richardson, S. and Tuna, I. (2003). 'Why are earnings kinky? An examination of the earnings management explanation'. *Review of Accounting Studies*, 8(2/3): 355–384.

Degeorge, F., Patel, J. and Zeckhauser, R. (1999). 'Earnings management to exceed thresholds'. *Journal of Business*, 72 (1): 1–33.

Dharmadhikary, S. and Joag-Dev, K. (1988). *Unimodality, Convexity and Applications*, Academic Press.

Durtschi, C. and Easton, P. (2005). 'Earnings management? The shapes of the frequency distributions of earnings metrics are not evidence ipso facto', *Journal of Accounting Research*, 43(4): 557–592.

Easton, P. and Sommers, G. (2003). 'Scale and the scale effect in market-based accounting research'. *Journal of Business, Finance & Accounting*, 32(1/2): 25–56.

Ellerman, D. (1985). 'The mathematics of double entry bookkeeping'. *Mathematics Magazine*, 58: 226–233.

Fama, E. and French, K. (2000). 'Forecasting profitability and earnings'. *Journal of Business*, 73(2): 161–175.

Frecka, T. and Hopwood, W. (1983). 'The effect of outliers on the cross-sectional distributional properties of financial ratios'. *Accounting Review*, 58(1): 115–128.

Flynn, M. (2006). 'Fitting human exposure data with the Johnson S_B distribution'. *Journal of Exposure and Environmental Epidemiology*, 16(4): 56–62.

Gore, P., Pope, P. and Singh, A. (2007). 'Earnings management and the distribution of earnings relative to targets: UK evidence'. *Accounting and Business Research*, 37(2): 123–150.

Grambovas, A., Giner, B. and Christodoulou, D. (2006). 'Earnings conservatism: Panel data evidence from the EU and the US'. *Abacus*, 42(3): 354–378.

Härdle, W. (1990). *Applied Nonparametric Regression*. Econometric Society Monographs, Cambridge University Press.

Heo, J., Kho, Y. and Shin, H. (2008). 'Regression equations of probability plot correlation coefficient test statistics from several probability distributions'. *Journal of Hydrology*, 355(1–4): 1–15.

Ijiri, Y. (1965). 'Management goals and accounting for control', Vol. 3 in Theil, H. *Studies in Mathematical and Managerial Economics.* North-Holland Publishing Company.

Ijiri, Y. and Simon, H. (1977). *Skew Distributions and the Sizes of Business Firms.* North-Holland Publishing Company.

Ioannides C., Peel, D. and Peel, M. (2003). 'The time series properties of financial ratios: Lev revisited'. *Journal of Business Finance and Accounting*, 30(5/6): 699–714.

Jacob, J. and Jorgensen, B. (2007). 'Earnings management and accounting income aggregation'. *Journal of Accounting & Economics*, 43(2–3): 369–390.

Johnson, N. (1949). 'Systems of frequency curves generated by methods of translation'. *Biometrika*, 36(1/2): 149–176.

Kahneman, D. and Tversky, A. (1979). 'Prospect theory: An analysis of decision under risk'. *Econometrica*, 2: 263–292.

Kerstein, J. and Rai, A. (2007). 'Intra-year shifts in the earnings distribution and their implications for earnings management'. *Journal of Accounting and Economics*, 44 (3): 399–419.

Kottegoda, N. (1987). 'Fitting Johnson S_B curves by the method of maximum likelihood to annual maximum daily rainfalls'. *Water Resources Research*, 23(4): 728–732.

Kotz, S. and Van Dorp, J. (2004). Beyond Beta: Other Continuous Families of Distributions with Bounded Support and Applications. World Scientific Publishing.

Lau, H., Lau, A. and Gribbin, D. (1995). 'On modelling cross sectional distributions of financial ratios'. *Journal of Business Finance and Accounting*, 22(4): 521–549.

Leuz, C., Nanda, D. and Wysocki, P. (2003). 'Earnings management and investor protection: an international comparison'. *Journal of Financial Economics*, 69(3): 505–527.

Levene, H. (1960). 'Robust tests for equality of variances'. In I. Olkin et al. (eds.), *Contributions of Probability and Statistics: Essays in Honor of Harold Hotelling*, Stanford University Press: 278–292.

Mage, D. (1980). 'An explicit solution for S_B parameters using four percentile points'. *Technometrics*, 22(2): 247–251.

Maijoor, S. and Vanstraelen, A. (2006). 'Earnings management within Europe: The effects of member state audit environment, audit firm quality and international capital markets'. *Accounting and Business Research*, 36(1): 33–52.

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

McLeay, S. (1997). 'Boundary conditions for ratios with positively distributed components'. *Journal of Business, Finance & Accounting*, 24(1): 67–83.

McLeay, S. and Trigueiros, D. (2002). 'Proportionate growth and the theoretical foundations of financial ratios'. *Abacus*, 38(3): 297–316; and 'A clarification', (2003) 39: v.

McLeay, S. and Stevenson, M. (2009). 'Modelling the longitudinal properties of financial ratios'. *Applied Financial Economics*, 19(4): 305–318.

Michalewicz, Z. and Fogel, D. (2002). *How to Solve It: Modern Heuristics*, Springer-Verlag.

Müller, H. (1984). 'Smooth optimum kernels estimators of densities, regression curves and modes'. *Annals of Statistics*, 12(2): 766–774.

Pagan, A. and Ullah, A. (1999). *Nonparametric Econometrics*. Cambridge University Press.

Parzen, E. (1962). 'On estimation of a probability density function and mode'. *The Annals of Mathematical Statistics*, 32(3): 1065–1076.

Peasnell, K., Pope, P. and Young, S. (2005). 'Board monitoring and earnings management: Do outside investors influence abnormal accruals?'. *Journal of Business, Finance & Accounting*, 32(7/8): 1311–1346.

Peel D., Peel M. and Venetis, I. (2004), 'Further empirical analysis of the time series properties of financial ratios based on a panel data approach'. *Applied Financial Economics*, 14(3): 155–163.

Petrovits, C. (2006). 'Corporate-sponsored foundations and earnings management'. *Journal of Accounting and Economics*, 41(3): 335–362.

Prais, S. (1976). *The Evolution of Giant Firms in Britain a Study of the Growth of Concentration in Manufacturing Industry in Britain: 1909–70*, No. 30 in Economic and Social Studies, Cambridge Press.

Rosenblatt, M. (1956). 'Remarks on some nonparametric estimates of a density function'. *Annals of Mathematical Statistics*, 27(3): 832–837.

Salgado-Ugarte, I., Shimizu, M. and Taniuchi, T. (1995). 'Practical rules for bandwidth selection in univariate density estimation'. *Stata Technical Bulletin Reprints*, 5: 5–19.

Siekierski, K. (1992). 'Comparison and evaluation of three methods of estimation of the Johnson S_B distribution'. *Biometrical Journal*, 34(7): 879–895.

Silverman, B. (1986). Density Estimation for Statistics and Data Analysis. Chapman and Hall.

Singh, A. and Whittington, G. (1968). *Growth, Profitability and Valuation: A Study of United Kingdom Quoted Companies*. Occasional Papers No. 7, Department of Applied Economics, University of Cambridge, Cambridge University Press.

Swain, J., Venkatraman, S. and Wilson, J. (1988). 'Least squares estimation of distribution functions in Johnson's translation system'. *Journal of Statistical Computation and Simulation*, 29: 271–297.

Tippett, M. (1990). 'An induced theory of financial ratios'. *Accounting and Business Research* 21(81): 77–85.

Tippett, M. and Whittington, G. (1995). 'An empirical evaluation of an induced theory of financial ratios'. *Accounting and Business Research*, 25(3): 208–218.

Trigueiros, D. (1995). 'Accounting identities and the distribution of ratios'. *British Accounting Review,* 27(2): 109–126.

Trigueiros, D. (1997). 'Non-proportionality in ratios: an alternative approach'. *British Accounting Review,* 29(2): 213–230.

Tsionas, E. (2001). 'Likelihood and posterior shapes in Johnson's S_B system'. *Sankhyā: The Indian Journal of Statistics*, 63(1): 3–9.

Vogel, R. (1986). The probability plot correlation-coefficient test for the normal, lognormal, and Gumbel distributional hypotheses. *Water Resources Research*, 22(4): 587–590.

Willett R. (1991). 'Transactions theory, stochastic processes and derived accounting measurement'. *Abacus*, 27(2): 117–134.

Whittington, G. and Tippett, M. (1999). 'The components of accounting ratios as co-integrated variables'. *Journal of Business Finance and Accounting*, 26(9/10): 1245–1273.

Yoon, S. (2005). 'A comparison of earnings management between KSE firms and KOSDAQ firms'. *Journal of Business, Finance & Accounting*, 32(7/8): 1347–1372.

Zhou, B. and McTague, J. (1996). 'Comparison and evaluation of five methods of estimation of the Johnson system parameters'. *Canadian Journal of Forest Research*, 26(6): 928–935.