Malaysian Management Journal 1 (2), 59 - 66 (1995)

The Use of Logarithmic Transformation and Numerical Derivatives for Business Performance Data

MARVIN D. TROUTT, SURESH K. TADISINA & HAROLD K. WILSON
Department of Management, Southern Illinois University, Carbondale, USA.

Key Words and Phrases: Graphics, Log Transformation, Relative Rates of Change, Numerical Derivatives, Fourth Generation Software, Decision Support Systems.

ABSTRAK

Nilai perlakaran data pencapaian perniagaan terhadap masa pada graf logaritma-separa pernah diketengahkan, tetapi dengan cara sambil lewa sahaja. Walaupun beberapa contoh dibincang, sokongan bermatematik tidak diberikan. Dalam kertas ini, kami sampaikan sokongan sedemikian untuk memperkenalkan kepada pembaca kaedah bergraf yang mudah tetapi amat berguna ini. Kami juga membincang dan menghuraikan suatu lanjutan idea tentang teknik ini. Kaedah ini khususnya berguna dalam pakej "spreadsheet" yang mempunyai keupayaan graf.

ABSTRACT

The value of plotting business performance data over time on a semi-logarithmic graph has been advocated, but in a somewhat off-hand and intuitive way. While examples have been discussed, no mathematical justification has been given. In this paper, we give such a justification in order to underscore, and perhaps acquaint the reader with this simple yet powerful graphics tool. We also discuss and illustrate an extension of the idea underlying this technique. These methods are particularly useful with spreadsheet packages which include graphics capabilities.

INTRODUCTION

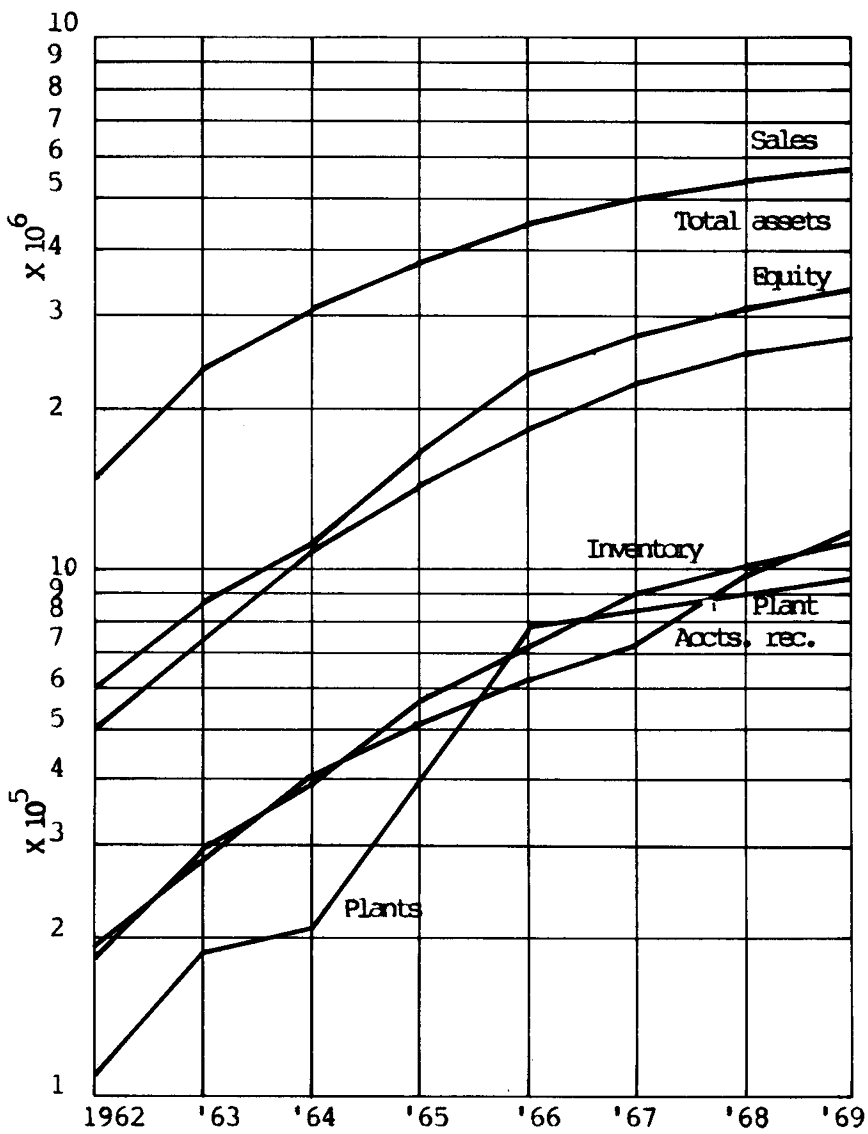
Buskirk (1971) has argued that plotting performance data over time on a semi-logarithmic graph as in Figure 1 provides two significant advantages. The first and most obvious is that trends become immediately apparent as against comparing the data at perhaps just two distinct points of time. More critically, however, he argues that time rates of change, as opposed to levels of the variables, can then be easily judged. In addition, he claims that analysis can then determine, for example, whether inventory or accounts receivable have been increasing more rapidly than sales volume. Thus, such a graph is argued to actually indicate the relative rates of change of those variables.

While the case for analysing relative rates of change is well supported in Buskirk (1971), the point that the semi-logarithm graph does, in fact, enable this analysis was not directly justified. Hence, our first purpose here is to provide such justification. We also discuss some limitations and extensions of this concept.

It is convenient to call attention to the fact that if a time series X_t is plotted on a (vertical log) semi-logarithm graph, the result is equivalent to plotting log X_t on a regular graph. This observation is of practical importance since fourth generation software such as LOTUS 123, SAS, FOCUS, IFPS and other decision support system generator products provide built-in logarithm functions along with graphics capabilities.

Also, while Buskirk (1971) proposed by implication the use of logarithms with base of 10, we argue that use of natural logarithms (i.e., with base of e) is superior. We denote the

FIGURE 1. Semi-logarithmic graph of performance data over time



natural logarithm function by ln(.) below. From the point of view of relative rates of change, either type of logarithm is suitable. Most fourth generation software provide both the base 10 and natural logarithm functions. However, the basic formulae are simplified if natural logarithms are used. Graphic displays will be largely unaffected as one is a constant multiple of the other. Hence, graphics using either logarithm base will be of the same shape.

With the background now developed, we give a mathematical rationale for the above claim in the next section.

RELATIVE RATES OF CHANGE

While time series data on balance sheet, and profit and loss data are collected discretely over time, it is conceptually advantageous to regard these as discretely sampled values of a continuous function of time, whether these be themselves inherently rates (e.g., sales) or levels (e.g., inventory).

For example, suppose $x(t)$ represents inventory level at time t. Suppose also that $s(t)$ indicates annual sales at time t. For the continuous function setting, $s(t)$ is interpreted as total sales for the year ending at time point t. If we graph the natural logarithm of $x(t)$ and $s(t)$, we are dealing with the functions :

$$X(t) = \ln x(t) \text{ and} \quad (1)$$

$$S(t) = \ln s(t) \quad (2)$$

From calculus, notice that the derivatives of $X(t)$ and $S(t)$ are given respectively by:

$$X'(t) = x'(t) \quad \text{and} \quad (3)$$

$$S'(t) = s'(t) \quad (4)$$

$$\frac{1}{x(t)}$$

$$\frac{1}{s(t)}$$

Hence, the contribution of the log graph is now clear. Namely, the rate of change of the log graph is that of the original function as a fraction of its level. Thus, while $x'(t)$ and $s'(t)$ may not be comparable, $X'(t)$ and $S'(t)$ are indeed comparable. A numerical example will help clarify the distinction.

Example 1:
Suppose at a given time point, we have $s(t)$ = RM200,000 per year and $x(t)$ = RM50,000. Suppose also $s'(t)$ = RM15,000 and $x'(t)$ = RM10,000. Hence, in absolute numerical terms, $s(t)$ is increasing faster than $x(t)$. However, note that $s'(t)/s(t) = .075$ while $x'(t)/x(t) = .20$. Thus, as a fraction of their respective levels, $x(t)$ is increasing much faster than sales.

Hence, the rates of change of $\ln s(t)$ and $\ln x(t)$ are comparable, providing a justification for Buskirk's suggestion. In the next section, we provide a caution and suggest an improved approach for some variables.

A CAUTION FOR SOME VARIABLES

While taking logarithms as discussed above will put all variables on the same percentage basis with respect to rate of change (graphical slope), this may be somewhat misleading for certain classes of variables. In particular, we may distinguish the two broad classes of variables consisting of control or independent variables and response or criterion variables. Typically, the latter are nonlinear functions of the former. For example, consider the familiar Cost-Volume-Profit analysis model:

$$\begin{aligned} p(t) &= \text{price} \\ r(t) &= \text{gross} \\ c(t) &= \text{revenue} \\ s(t) &= \text{sales} \end{aligned} \quad (5)$$

$$\text{for which } r(t) = p(t)\,s(t) - c(t) \quad (6)$$

The reader may easily check that the log of $r(t)$ is a complex function of the independent variables of the model. In particular, rates of change, relative or not, of the $p(t)$, $s(t)$, and $c(t)$ values are not readily comparable to that of $r(t)$. In fact,

$$r'(t) = p(t)\,s'(t) + p'(t)\,s(t) - c'(t) \quad (7)$$

For variables of this type, it may be safer to plot actual derivatives if the goal is to judge trend relationships. The techniques for finding the derivatives of a numerical discrete series will now be illustrated. The reader will note that the procedure is very amenable to spreadsheet modelling.

We now wish to show how to obtain the value of r'(t) and also r''(t) in equation (7) above. Since we must work with sampled values of the time series for time points before a given current time, we need numerical formulae in terms of backward differences of the series.

Following Kellison (1975), we define:

$$\nabla s(t) = \text{First Backward Difference of } s(t) = s(t) - s(t-1) \qquad (8)$$

When used with the power notation, $\nabla^n s(t)$ means n successive applications of the ∇ operation. For instance:

$$\begin{aligned}\nabla^2 s(t) &= \nabla\ [\nabla\ s(t)] \\ &= \nabla\ [s(t) - s(t-1)] \\ &= s(t) - 2s(t-1) + s(t-2) \qquad (9)\end{aligned}$$

Now from Kellison (1975) we have, for the derivative of s(t) at time t, denoted s'(t),

$$s'(t) = s(t) + (1/2\nabla^2 s(t) + (1/3)\nabla^3 s(t) + \ldots \qquad (10)$$

It will be useful also to note that the second derivative of s(t) is given by the series:

$$s''(t) = \nabla^2 s(t) + \nabla^3 s(t) + (11/12)\nabla^4 s(t) + \ldots \qquad (11)$$

An example will now be given to clarify the formulae.

Example 2:

Suppose we have observed values of s(t) for t = 1,2,3,4, and 5 as follows:

t	s(t)
1	102
2	110
3	130
4	175
5	203

Find approximations to the first derivative, and the second derivative of s(t) at time t = 5.

The work can be conveniently put into the following spreadsheet model:

t	s(t)	∇s(t)	∇^2s(t)	∇^3s(t)	∇^4s(t)
1	102	.	.	.	.
2	110	8	.	.	.
3	130	20	12	.	.
4	175	25	5	-7	.
5	203	28	3	-2	5

Note that the backward difference at t = 1 is missing since we do not know s(0). Now plugging into formulae (10) and (11) as far as we have data (and also assuming higher order differences are zero) gives:

$$\begin{aligned}s'(5) &= 28 + 1/2(3) \text{ x } 1/3(-2) + 1/4(5) \\ &= 30\ 1/12 \text{ and} \\ s''(5) &= 3 + (-2) + 11/12(5) \\ &= 5\ 7/12\end{aligned}$$

In the next section, we discuss the practical importance of these approximations.

DERIVATIVES AS LEADING INDICATORS OF TROUBLE

In this section, we show how derivatives can be used with performance data to provide early indications of impending trouble. From elementary calculus it is known that when the derivative is positive, the function is increasing. Also, while the function may be increasing, its own rate of increase (measured by its second derivative) may itself be decreasing (i.e., negative). Hence, a negative second derivative can be used as an early indication of future decreases in a function such as gross revenue (6).

Specifically, we wish to push on beyond (7) to the second derivative of r(t) which will be denoted r''(t). From the rules of differentiation we obtain by differentiating (7) with respect to t again,

$$r''(t) = p(t)s''(t) + 2p'(t)s'(t) + p''(t)s(t) - c''(t) \tag{12}$$

Example 3 shows how (12) may be used in an early warning context.

Example 3:

Suppose that during a given period, prices have been constant at p(t) = 100 which implies that both p'(t) and p''(t) are zero. Suppose that s(t) = 100, s'(t) = 2, and s''(t) = -9 and that c''(t) = -3. Evaluating (12) shows that r''(t) = - 897. Hence, if present trends continue, gross revenue will eventually begin to decline, even though prices are constant, costs are improving and sales are increasing at the current time. Thus, r''(t) provides an early indication of impending declines in revenue, *ceteris parobus.*

ILLUSTRATION

The idea of employing approximate derivatives using differences and natural log transformations discussed above will be illustrated using sales and net income before tax (NIBT) data pertaining to Ford Motor Co. for the time period 1974 - 1983.

Examining the sales data plotted in Figure 2, an increasing trend is observed up to 1979 followed by a decline/marginal increase during the next three years and finally a significant increase in 1983. This pattern can be easily seen by examining the first differences (represented as spikes in Figure 2). Further, the plot of first differences also indicates the pattern of changes in sales — for example, the increase in sales was increasing over the 1975 - 77 period; however, it decreased over the next three years and, in fact, the improvement was negative in 1981. Therefore, the decrease in sales increase in 1978 would be an early warning signal indicating a potential problem situation.

Similarly, if the NIBT data were examined (see Figure 3), the plot would indicate an increasing trend up to 1977 followed by a decreasing trend up to 1980 and an increasing trend from then on. The first differences plot conveys this information very explicitly and further indicates that in 1976 and 1977 the increase in NIBT over the previous year was about the same and started going negative in 1978.

Using the first difference plots, one can anticipate problems in the performance measures or, more generally, predict turning points in the time series.

Viewing the sales and NIBT plots together in one graph (see Figure 4) there is an indication that although sales revenue increased in 1978 and 1979, NIBT started declining in 1978 itself. Examining the plots of the same data after log transformations (see Figure 5), the relative rates of change in the performance measures become comparable. For example, in 1975 sales relative to previous year increased marginally whereas NIBT relative to previous year decreased substantially, and in 1976, the relative change in NIBT over the previous year was far more significant than the relative change in sales over the previous year. Therefore, the use of log transformations facilitates the comparison of relative changes in various performance measures in one plot whereas such comparisons would not be possible using raw data plots.

CONCLUSION

The natural log transformation of performance data time series places each series on a comparable basis which allows direct comparison of relative rates of change. For series which are functions of other series, relative rates of change may not be suitable. In such cases, actual first and second derivatives may be more informative. In fact, the second derivative serves as an early warning of impending trouble in such series. Numerical methods for finding these derivatives in a spreadsheet model are also given. Finally, the advantages of using log transformations and numerical derivatives are illustrated with an example. In effect, the paper emphasises the

FIGURE 2. Sales data and first differences plot

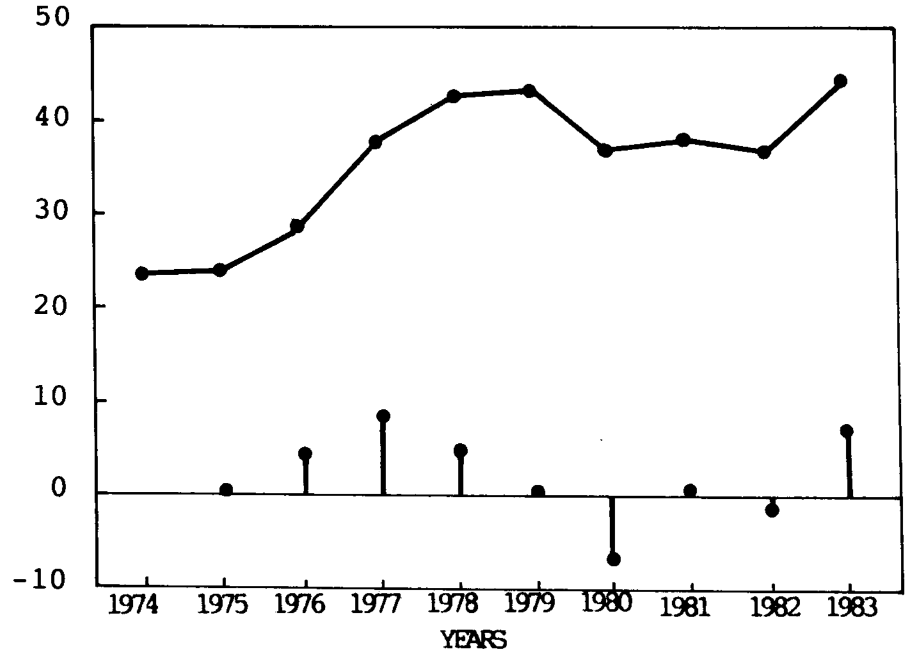


FIGURE 3. NIBT data and first differences plot

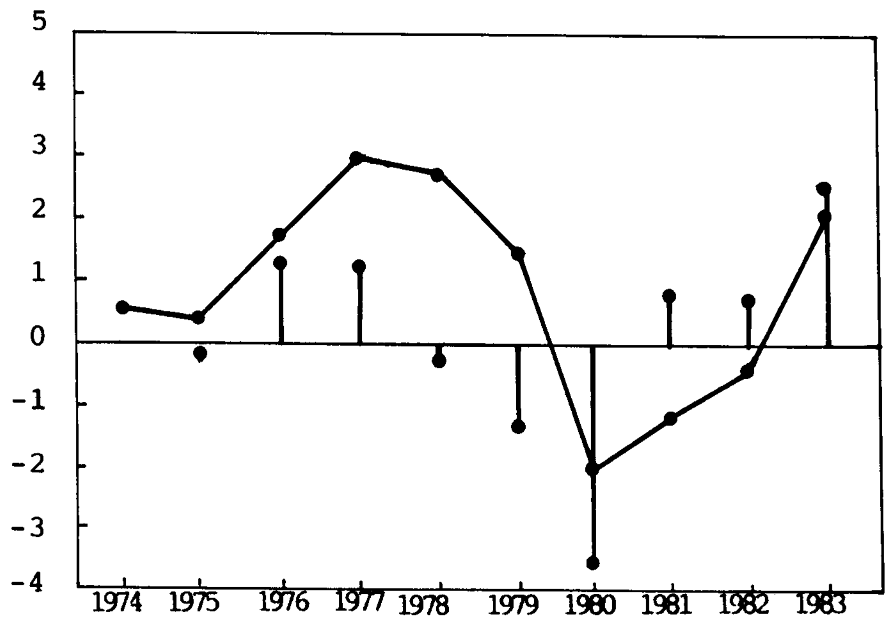


FIGURE 4. Sales and NIBT data plot

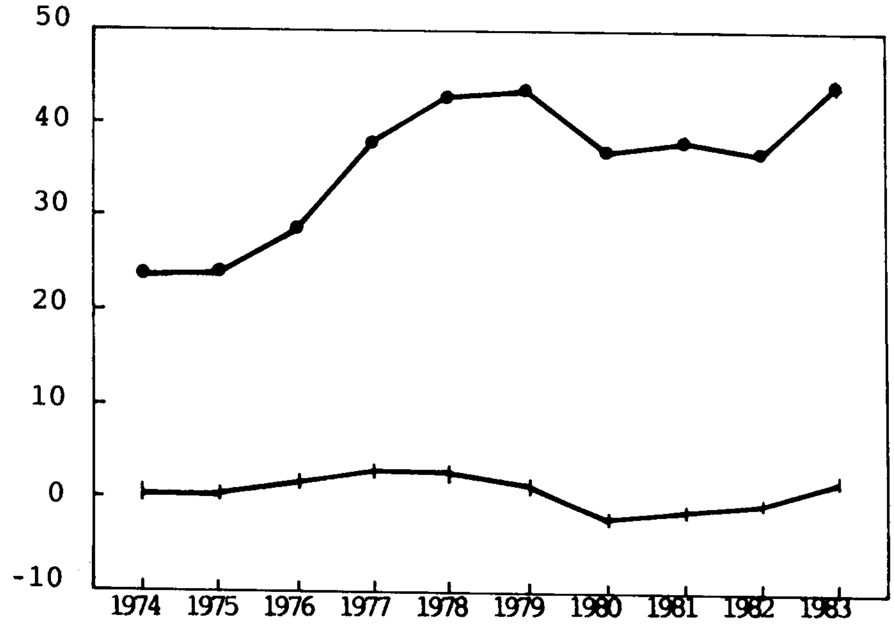
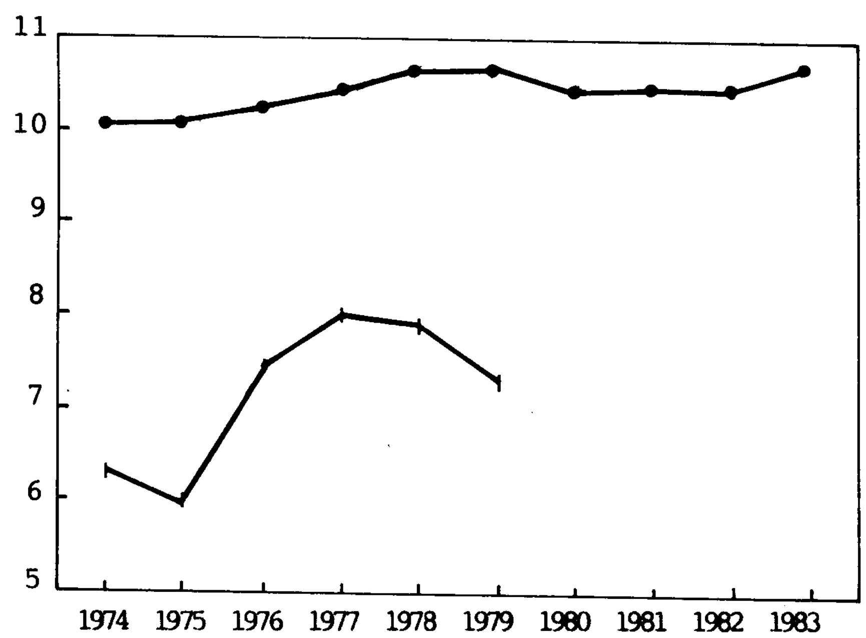


FIGURE 5. Natural log transformed sales and NIBT* data plot



NIBT for 1980, 1981, and 1982 was negative and therefore the data points for the three years do not appear in the plot, as the natural log transformations for negative values are not defined.

usefulness of plotting log transformed data and numerical derivates for comparative purposes, and for viewing trends and identifying turning points in analysing business performance data.

REFERENCES

Buskirk, R.H. (1971). *Business and Administrative Policy : Text, Cases, Incidents, and Readings,* New York : John Wiley & Sons, Inc..

Kellison, S.G. (1975). *Fundamentals of Numerical Analysis,* Illinois : Richard D. Irwin, Inc..