

THE VARIANCE OF FIRM GROWTH RATES: THE 'SCALING' PUZZLE

John SUTTON
London School of Economics

Certain recently reported statistical regularities relating to the dispersion of firms' growth rates have begun to attract attention among IO economists. These relationships take the form of power law or scaling relationships and this has led to suggestions that the underlying mechanisms which drive these relationships may have some interesting analogies with the mechanisms which drive scaling relationships in other fields.

In the first half of this talk, I report some new empirical evidence in this area and I put forward a new candidate explanation for the relationships we observe. This part of the talk is based upon the attached paper.

Notwithstanding the rather striking and clear cut empirical evidence that has been reported in the literature for the existence of such relationships, there appears to be some skepticism among economists as to whether these relationships are of any substantial interest. In the second half of this talk, I will focus on presenting an extended application in the area of industry dynamics, in which the uncovering of the appropriate size variance relationship leads to the possibility of a very simple and useful characterization, which in turn permits some new light to be thrown on the 'persistence of leadership' debate.

The Variance of Firm Growth Rates: The ‘Scaling’ Puzzle

John SUTTON

London School of Economics

1. INTRODUCTION

Ever since the pioneering work of Gibrat (1931), economists have taken an interest in the relationship between the size of a firm and its expected proportional growth rate (for a review, see Sutton 1996). Only recently, however has any systematic work been devoted to the related question which links firm size to the variance of growth rates: taking the set of firms whose annual sales revenue lies in the interval $(y, y+\epsilon)$, consider the variance σ^2 of the change in sales Δy from one year to the next. What is the relationship between $\sigma^2(\Delta y)$ and y ?

This relationship has been investigated in the recent physics literature, following Stanley et al. (1996)¹, who report a ‘power-law’ relationship of the form $\sigma^2(\Delta y) = Ay^x$. Stanley et al. write this relation in terms of the proportional growth rate $g \equiv \Delta y/y$, viz. $\sigma(g) \equiv \sigma(\Delta y/y) = \sigma(\Delta y)/y = Ay^r$ where $r = (x/2)-1$, and they (i) establish the ‘power-law’ form of the relation by showing that a plot of $\ln(\sigma/y)$ against $\ln y$ displays linearity over a wide range of y (a factor of about 10^6); and (ii) they estimate the coefficient r at -0.15 , corresponding to a value of $x = 1.7$. More recently, it has been shown that a similar empirical relationship holds not only for firms, but for national economies (Lee et al. (1998)).

Stanley et al. suggest that a natural model to have in mind here is one in which each firm consists of a number of equal sized units (businesses², divisions etc.), and they distinguish between the possibility that these units growth rates are statistically independent, and the

¹ Stanley et al. also investigate the shape of the distribution of proportional growth rates, which they find to be well represented by a double exponential distribution. This aspect of their findings lies beyond the scope of the present paper. See also Amaral et al. (1988), Pierrou et al. (1999).

² In what follows, the term ‘business’ is used to denote the activities of a firm within a single market; thus a firm comprises one or more businesses.

possibility that they are correlated in some way. The ‘independence’ case leads to a power law, but the coefficient r takes the value -0.5 , far below the reported empirical value of -0.15 .

The appearance of a power law (or ‘scaling’) relationship suggests an analogy with areas of statistical physics where such laws arise because of subtle forms of correlation between the units that comprise a system; this has raised the question of whether some similar forms of correlation might be at work in these economic settings. Following this idea, Stanley et al. (1996) proposed one specific candidate model of decision making in firms that might lead to correlated changes in sales by the constituent businesses. This type of candidate explanation raises a number of difficulties:

- (a) While we would expect that certain factors impinging on the firm as a whole might affect all its constituent businesses in a similar way, the large majority of factors affecting the sales of a business from one year to the next are of a kind that are specific to the market in which that business operates, and these would not be expected to affect the sales of the firm’s other businesses. (This claim is directly testable; see Section 2 below.) Any such firm level effects will make the $\ln \sigma(g)$, $\ln y$ relation ‘flatter’, relative to a model in which the growth rates of the constituent businesses are independent, but these effects, insofar as they are present, are probably small, and their magnitude is likely to fluctuate from year to year with the changing macroeconomic environment.
- (b) While it is possible to suggest various rather context-specific models of correlations, following Stanley et al. (1996) or otherwise, it is not clear why such a mechanism would lead to a power-law relationship. Moreover, the fact that a power-law relation holds for the very different context of national economies suggests that a relatively primitive and robust mechanism may be driving this result.

The argument that follows is developed in three steps:

- (i) A re-examination of evidence from the database used by Stanley et al. (1996) indicates that the value of the power coefficient r fluctuates across years, while remaining in the range -0.15 to -0.21 . Moreover, a direct test of whether growth rates of constituent businesses within a firm are correlated indicates that, while a positive correlation appears to be present, it is extremely weak and unstable.
- (ii) This suggests an exploration of the appropriate way to model the firm as a set of businesses, with a view to asking whether a simple benchmark model based on independent businesses might provide a useful (lower) bound to the observed values of r .

The most salient feature of this situation is that businesses vary widely in size; and larger businesses must, as a matter of logic, form parts of larger rather than smaller firms. Yet it is not clear how any ‘general’ relationship might hold between the size of a firm and the size distribution of its constituent businesses. The central idea in what follows lies in proposing a simple model, in which a firm with annual sales revenue y is represented as consisting of a set of businesses with sales $y_1, y_2 \dots y_i \dots$ with $\sum_i y_i = y$. We represent the y and the y_i s as integers, so that (y_i) is a partition of the integer y (Andrews (1970)). We set out to explore the consequences of hypothesising that, for a randomly selected firm of size y , all partitions of y are equally likely. Assuming that each business is subject to the same distribution of proportional shocks, it is shown that the size-variance relation for firms follows an (approximate) power law, with a coefficient in the range $1.585 \geq x > 1.5$ or a value of r in the range $-0.21 \geq r > -0.25$ which corresponds well to the (lower) bound of observed values.

- (iii) It is suggested that the presence of small ‘firm level’ shocks, whose size fluctuates across firms and between years, may explain the small residual fluctuations in the observed slope coefficient above this benchmark value.

2. RE-EXAMINING THE EVIDENCE³

The relationship was first re-examined for firm level data using the Compustat database for 1980-98, following the procedure used in Stanley et al. (1996) in which firms are grouped by size into 14 ‘bins’, the boundaries between successive bins being a factor of 4. The results⁴ indicate that the slope coefficient fluctuates in the range $-0.21 \leq r \leq -0.15$. Linearity is preserved in most years (Figure 1a) but in three years a substantial departure from linearity arises at the extreme ends of the distribution. These years are ones in which r is relatively low in absolute value and in all cases the nature of the departure is of the form illustrated for 1997 in Figure 1b. (It is interesting that the data illustrated in Figure 2 of Stanley et al. (1996) shows this same kind of departure from linearity at the extremes.) In view of the relatively small samples encountered for very large and very small firm sizes in this database it seems appropriate to draw no more precise conclusions on the basis of this evidence⁵.

For each of the years 1992-98, the Compustat database contains, for over 800 firms, data for several ‘segments’ of the firm⁶. Taking all firms for which at least two segments are reported, for each year from 1993 to 1997, a correlation coefficient was calculated between the rate of growth of the largest segment, and the rate of growth of the second largest segment, from that year to the next (Table 1 and Figure 2). The values fluctuate widely and the value for the pooled sample is very low (0.115). This suggests that, while some positive correlation (possibly reflecting ‘firm level effects’ on the sales of constituent businesses) may be present, such effects appear to be too small and too unstable to account for the observed linearity and flatness of the $\ln \sigma(g)$ versus $\ln y$ relationship. This observation motivates the investigation which follows.

³ I am grateful to Jian Tong for his assistance with the computations reported in this section.

⁴ The values for each year are illustrated in Figure 3 of Section 3 below.

⁵ For example, in 1997 the number of observations in the 14 bins were as follows (lowest bin first): 6, 9, 40, 83, 246, 478, 937, 1337, 1412, 1085, 587, 240, 32, 2.

⁶These ‘segments’ are defined by the firms themselves in their annual reports and may correspond either to individual businesses in the present sense, or to groups of businesses.

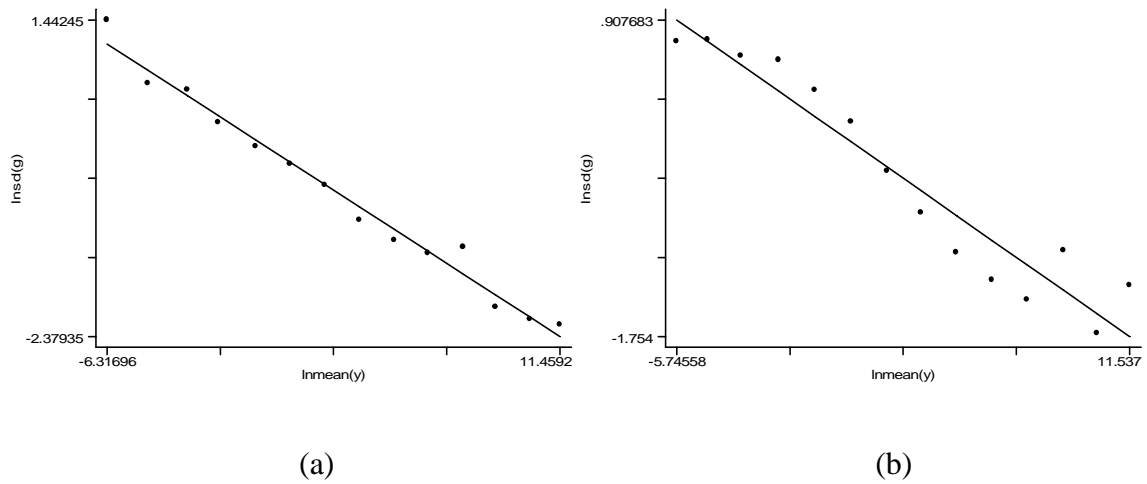


Figure 1. The values of $\ln \sigma(g)$ versus $\ln y$ for each bin, for 1993-94 (panel (a)) and 1997-98 (panel (b)). Following Stanley et al. (1996) firms are grouped into 14 bins, where the boundaries between successive bins differ by a factor of 4. The fitted regression lines are shown.

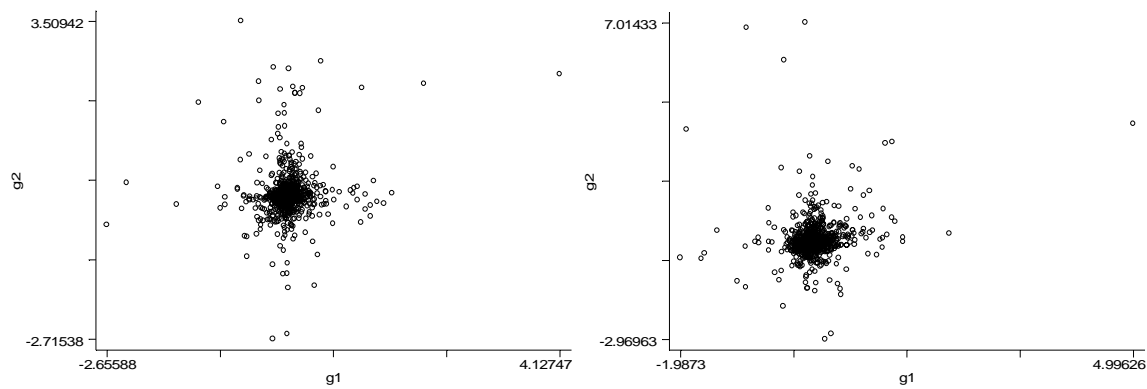


Figure 2. A scatter diagram showing the rate of growth of the second largest segment g_2 versus the rate of growth of the largest segment g_1 for each firm in 1993 (left hand panel) and 1995 (right hand panel).

year	'93-'94	'94-'95	'95-'96	'96-'97	'97-'98	pooled
Correlation Coefficient	.114	.062	.154	.079	.148	.115
No of Observations	955	970	1053	1073	868	4919

Table 1. Correlation coefficients between the year-to-year proportional growth rates of a firm's largest 'segment', and that of its second largest 'segment'.

3. A MODEL

Denote the size (annual sales revenue) of the firm by an integer n . Denote the sizes of its constituent businesses by integers n_1, n_2, \dots where $\sum_i n_i = n$, so that $\{n_i\}$ is a partition of n .

Assume⁷

- (i) for a randomly selected firm of size n , all partitions of n are equally likely.
- (ii) each constituent business experiences a proportional change in its size, denoted g , where g is an independent draw from some underlying distribution. To ease notation we normalize by setting the variance of the distribution of g to unity. Thus the change in size of a firm of initial size n has variance n^2 .

Assumption (ii) implies that the variance of the change in size of a firm of initial size n , whose constituent units are of sizes n_1, n_2, \dots is given by $n_1^2 + n_2^2 + \dots$. We aim, using Assumption (i), to average this value over all possible partitions of n .

The number of partitions of n is denoted $p(n)$; Table 1 illustrates the procedure for writing down the set of partitions, for successive integers. For example, if $n = 3$ the partitions are (1,1,1), (1,2) and (3). Here we have

$$\sigma^2(\Delta n) = \frac{1}{3}(1^2 + 1^2 + 1^2) + \frac{1}{3}(1^2 + 2^2) + \frac{1}{3}(3^2) = 17/3$$

In order to calculate the variance as a function of n , we first define a function $c(m,n)$, representing the number of occurrences of m in the set of partitions of n (Table 3).

⁷ Assumption (i) carries the main idea of the analysis. To facilitate comparisons with the model of 'equal sized units' mentioned earlier, the simple rule for growth rates of businesses embodied in Assumption (ii) is maintained for the moment. In Annex 2, this assumption is replaced by a more economically appropriate representation, but the results which follow remain unchanged.

$p(1) = 1$	1	1	1	1	1	1	1
$p(2) = 2$	1	1	1	1	1	2	
$p(3) = 3$	1	1	1	1	3		
	1	1	1	2	2		
$p(4) = 5$	1	1	1	4			
	1	1	2	3			
$p(5) = 7$	1	1	5				
	1	2	2	2			
	1	2	4				
	1	3	3				
$p(6) = 11$	1	6					
	2	2	3				
	2	5					
	3	4					
$p(7) = 15$	7						

Table 2. A table of partitions. For $n = 0$, $p(0) = 1$. Dotted lines show the table of partitions of each integer in turn, $1 \leq n \leq 7$. Note that 1 has only one partition, i.e. 1. For any $n \geq 2$, the table for n is constructed from the table for $n - 1$ using the following rule: adjoin 1 on the left to each partition of $n - 1$. Now take each partition of $n - 2$ whose least part is not less than 2; adjoin 2 on the left to each such partition. Continuing for $j = 3, 4, \dots, [n/2]$, take each partition of $n - j$ whose least part is not less than j ; adjoin j on the left of each such partition. This procedure ends with $j = [n/2]$; here, j is adjoined to the single part $n - [n/2]$. Finally, the single partition n is added.

n	m =						
	1	2	3	4	5	6	7
1	1						
2	2	1					
3	4	1	1				
4	7	3	1	1			
5	12	4	2	1	1		
6	19	8	4	2	1	1	
7	30	11	6	3	2	1	1

Table 3. A table of the function $c(m,n)$, showing the number of instances of m in the set of partitions of n , for $1 \leq n \leq 7$.

In this notation, on appealing to assumption (i), we have

$$\sigma^2(\Delta n) = \frac{1}{p(n)} \sum_{m=1}^n c(m, n) \cdot m^2$$

To proceed, we need an expression for $c(m, n)$. It can be shown that $c(m, n)$ can be represented in terms of the partition function $p(n)$ as follows⁸:

$$c(m, n) = \sum_{k=1}^{\lfloor n/m \rfloor} p(n - mk)$$

Using this relation, and inserting the value of $p(n)$, it is possible to compute $\sigma^2(\Delta n)$ for small values of n . Figure 3 shows a plot of $\ln \sigma^2(\Delta n)$ versus $\ln n$; it is clear that the relation is (approximately) represented as a ray through the origin, corresponding to a power-law relation of the form $\sigma^2(\Delta n) = A n^x$. (Under the normalisation chosen above, when $n = 1$ we write $\sigma^2(\Delta n) = 1$, so $A = 1$, whence $\ln \sigma^2(\Delta n) = \ln n = 0$.) The slight departure from a power-law relation can be illustrated by examining the value of $\ln \sigma^2(\Delta n) / \ln n$, corresponding to the slope of a ray from the origin to the n -th observation in Figure 3, for successive values of n (Table 4).

⁸ This follows from Theorem 3 of Fine (1980) on choosing Fine's sequence (a_j) as follows: $a_m = 1$, $a_j = 0$ for all $j \neq m$. I am grateful to George Andrews for this observation. Thus for $n = 7$, we have

$$\begin{array}{l} c(1,7) = p(6)+p(5)+p(4)+p(3)+p(2)+p(1)+p(0) \\ c(2,7) = \qquad \qquad p(5) \qquad \qquad +p(3) \qquad \qquad +p(1) \\ c(3,7) = \qquad \qquad \qquad p(4) \qquad \qquad \qquad +p(1) \end{array}$$

$$\begin{array}{l} \text{while } c(4,7) = p(3); c(5,7) = p(2); c(6,7) = p(1) \\ \text{and } c(7,7) = p(0) = 1. \end{array}$$

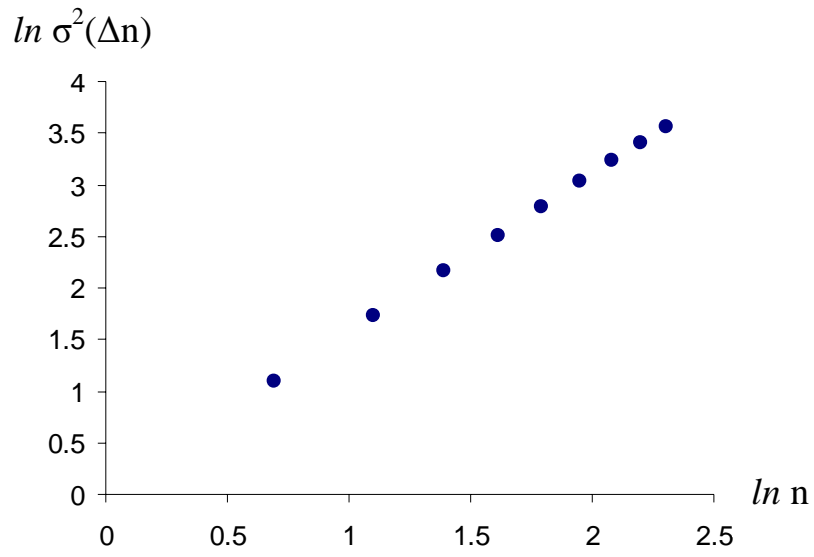


Figure 3. The predicted $\ln \sigma^2(\Delta n)$, $\ln n$ relation under the partitions model for $n = 1$ to 10.

n	$\ln(\sigma^2(\Delta n))/\ln n$	n	$\ln(\sigma^2(\Delta n))/\ln n$
2	1.5850	20	1.5415
3	1.5789	30	1.5372
4	1.5687	40	1.5345
5	1.5651	50	1.5327
6	1.5600	60	1.5314
7	1.5580	70	1.5303
8	1.5549	80	1.5295
9	1.5530	90	1.5287
10	1.5511	100	1.5281

Table 4. An illustration of the departures of the function $\sigma^2(\Delta n)$ from a simple power law. The table shows the value of the ratio $\ln \sigma^2(\Delta n)/\ln n$ for selected values of n .

Direct computation indicates that the slope of this ray falls very slowly with n . The limiting behaviour of the relation in the limit $n \rightarrow \infty$ can be investigated analytically. It is shown in the Appendix that as $n \rightarrow \infty$, the function $\sigma^2(\Delta n)$ becomes asymptotic to a power law relation An^x where $x = 1.5$.

The claim of this paper is that this simple ‘partitions’ model provides a reasonable bound, corresponding to the case in which ‘firm effects’ (or correlations across businesses) vanish. This is illustrated in Figure 4, where the empirical data for $\ln \sigma(g)$ vs. $\ln y$ is shown together with the candidate bound for $x = 1.55$, ($r = 0.225$) corresponding to the ‘empirically plausible’ value $n = 10$.

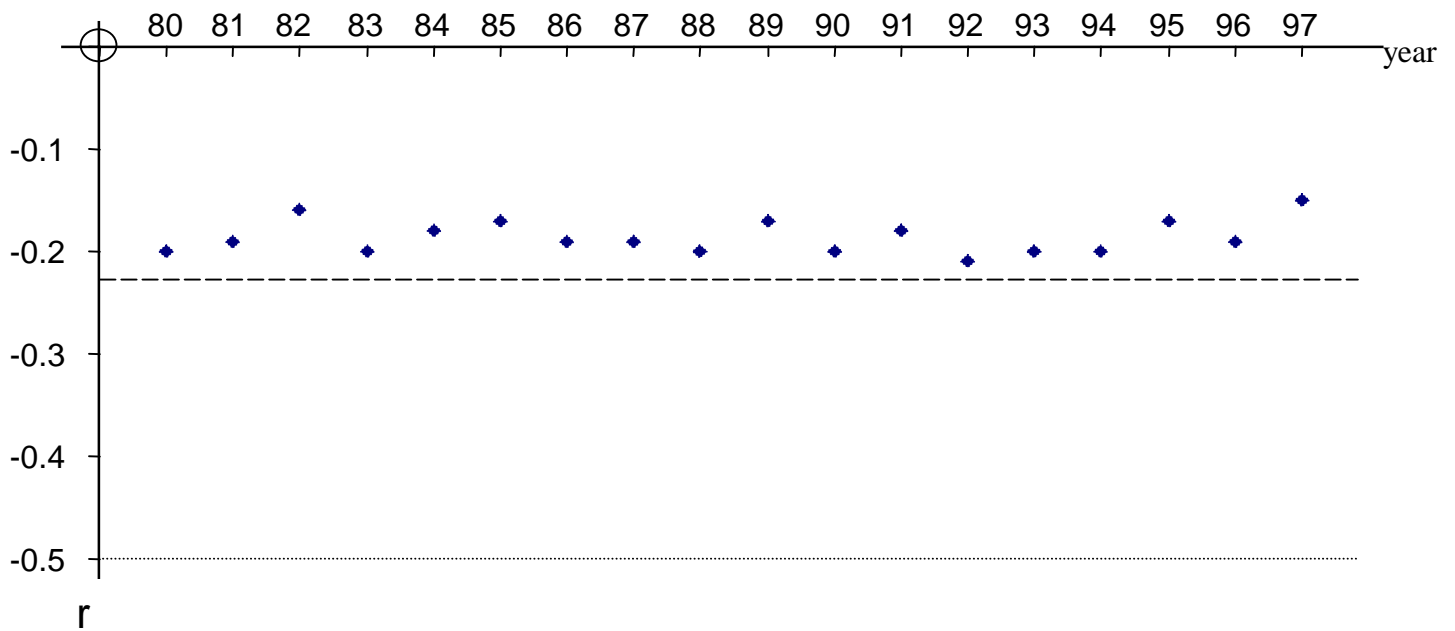


Figure 4. Empirical values of the slope coefficient r by year as estimated by the slope of the $\ln \sigma(g)$ versus $\ln y$ relation (Figure 1). The bound from the ‘partitions’ model of independent businesses is shown as a hatched line. The dotted line corresponds to the model in which each firm consists of independent units of a fixed size.

4. AN EXTENSION: RE-EXAMINING ASSUMPTION (ii)

The analysis in the preceding section begins by assuming that the firm's sales y may be divided into the sales $\{y_i\}$ of a set of businesses, and it proceeds to model the (integer valued) y_i as a partition of the (integer) y . From an economic point of view, this procedure might seem a little odd: it might seem more natural to think of the sales of each business, which by definition corresponds to the firm's sales in one particular market as being represented by a product $s_i Y_i$ where s_i is the firm's market share in market i , and Y_i is the total (i.e. industry) sales revenue of market i . Once y_i is represented in this way, it is natural to ask: within a single market i , what is the relation between a firm's size (as measured now by its market share in that market) and the variance of its change in size (market share) from one year to the next?

It is difficult to address this question using standard economic datasets: records are rarely available of the market shares of even a set of leading firms for a usefully long number of years. The following discussion draws on some results arising from a dataset for Japan, compiled by the author, which contains the market shares of a set of 'largest' firms within 50 well defined markets in each year from 1973 to 1997. For each market share band, the variance measure $\sigma^2(\Delta s)$ was computed. The data is illustrated in Figure 5. The 'size-variance of growth' relation was investigated by running a regression of $\ln(\sigma(\Delta s)/s)$ versus $\ln s$, together with industry dummies. The results are consistent with a power-law relation with slope $r \cong -0.5$ corresponding to the relation $\text{var}(\Delta s) = \text{constant} \cdot s$. This relation can be interpreted by reference to a model in which a firm's market share consists of its sales to a number of 'independent customers' of equal size. The slope coefficient r , in other words, corresponds to the reference model of Stanley et al. noted in Section 1 above. This begs the question: how can this result be reconciled with the much flatter slope coefficients ($|r| \leq 0.21$) found for firm level data?

The two results are easily reconciled by appealing to the idea that the market share s is drawn from some given distribution; but the size of the industry is drawn from a distribution which is very diffuse. This idea is made precise in what follows.

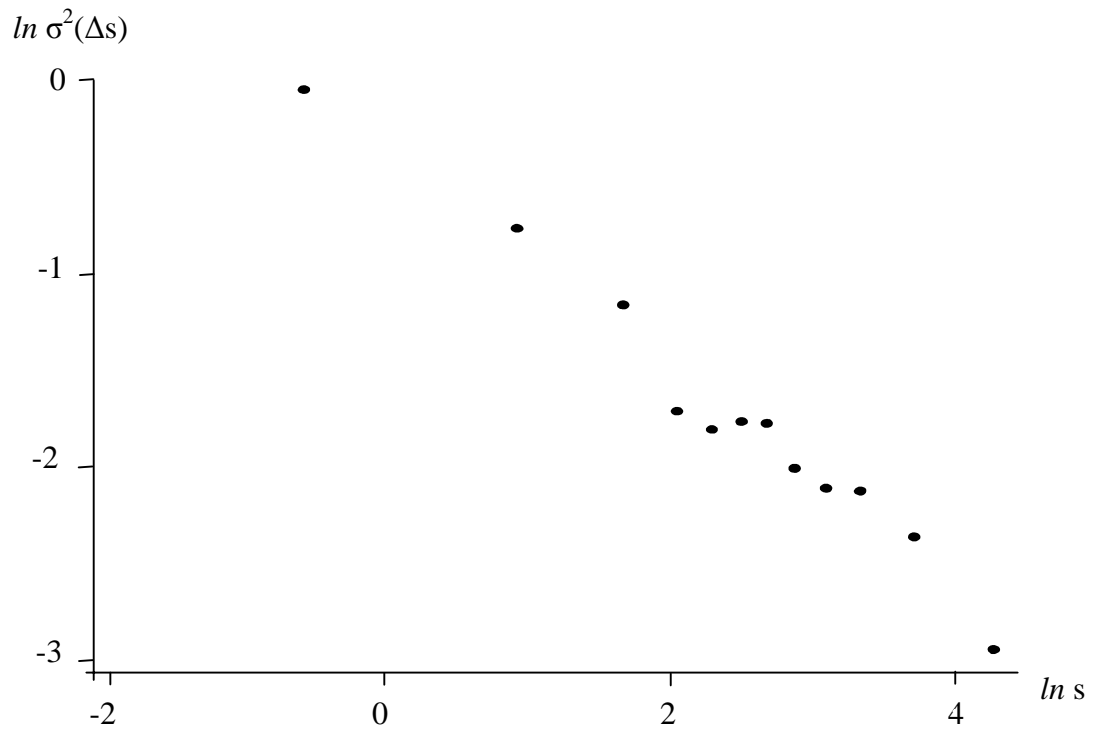


Figure 5. The $\ln \sigma^2(\Delta s)$ versus $\ln s$ relation for Japanese manufacturing firms. The data relates to a set of leading firms in each of 45 narrowly defined industries from 1973 to 1997. The pooled observations are split into twelve bins containing equal numbers of observations. A fitted regression incorporating industry dummies gives a slope of -0.5 .

Let the sales of the firm y be represented, as before, as the sum of the sales of a number of businesses in a series of markets indexed by i , so that the sales of these businesses, y_1, y_2, \dots form a partition of y , all partitions being equally likely. Now, however, interpret each y_i as being a product of the form $s_i Y_i$, viz.

$$y = \sum_i y_i = \sum_i s_i Y_i$$

In what follows, the sales of a single business will be denoted y , the subscript i being dropped to ease notation. Assume that shocks to y derive wholly from fluctuations in market share, the industry sales revenue Y being constant. It follows that

$$\text{var} (\Delta y) = Y^2 \text{var} (\Delta s)$$

Now we assume that s is a draw from some (discrete) distribution, whence it takes values $s_1, s_2, \dots, s_j, \dots$ with associated probabilities $p_1, p_2, \dots, p_j, \dots$ so that the sales revenue y corresponds with probability p_j to a market share of s_j in a market whose size Y equals y/s_j (In Bayesian terms, this amounts to treating the distribution of Y as a diffuse prior.).

It now follows that the variance of the change in sales of a business of size y is given by

$$\begin{aligned} \text{var} (\Delta y) &= Y^2 \text{var} (\Delta s) \\ &= \sum_j p_j \left(\frac{y}{s_j} \right)^2 \text{var} (\Delta s_j) \end{aligned}$$

Using the empirical result cited above to write $\text{var} (s_j) = \text{constant} \cdot s_j$, it follows that

$$\text{var} (\Delta y) = \text{constant} \cdot y^2 \sum_j \frac{p_j}{s_j}$$

Since $\sum(p_j/s_j)$ is a constant which depends only on the underlying distribution of s , it follows that $\text{var} (\Delta y)$ is proportional to y^2 , so that $\sigma(\Delta y/y) \equiv \sigma(g)$ is independent of y , whence Assumption (ii) above, and so the results of the basic model of Section 3, follow directly.

5. CONCLUSIONS

It has been suggested that the power-law nature of the $\sigma(g)$ versus y relationship reflects the workings of some form of correlation of growth rates across the firm's constituent businesses. Here, it has been shown that a simple model which allows for the fact that these businesses vary in size, as modelled by a simply 'partitions' of independent businesses, provides a good first order approximation to what is observed empirically. One candidate explanation for the fluctuations in the value of the slope from year to year is that these are associated with weak 'firm level' effects whose size fluctuates from one year to the next; whether or not this is so remains an open question. The relationship derived from the partitions model of independent businesses should be thought of as providing only a lower bound to the (negative) value of the slope coefficient r . The presence of correlations in the growth rates of a firm's constituent businesses will lead to a flattening of the $\ln \sigma(g)$ versus $\ln y$ relationship.

APPENDIX

Using the expressions for $\sigma^2(\Delta n)$ and $c(m,n)$ in Section 2, the expression for the variance can be written either as

$$\sum_{m=1}^n m^2 \left[\sum_{k=1}^{\lfloor n/m \rfloor} \frac{p(n-mk)}{p(n)} \right] \quad (1)$$

where the inner sum is taken within rows in Table 5, (and where $\lfloor \cdot \rfloor$ denotes the integer part) or by first summing within columns leading to

$$\sum_{m=1}^n s(m) \frac{p(n-m)}{p(n)} \quad (2)$$

where $s(m)$ is defined as the sum of squares of the divisors of m (whence $m^2 \leq s(m) < cm^2$ where c is a constant (≈ 1.6)).

Using the standard asymptotic formula

$$p(n) \approx \frac{1}{4\sqrt{3}} \frac{1}{n} e^{b\sqrt{n}} \quad \text{where } b = \pi\sqrt{\frac{2}{3}} \quad (3)$$

to write

$$\frac{p(n-m)}{p(n)} \approx \frac{n}{n-m} e^{-b(\sqrt{n}-\sqrt{n-m})} \quad (4)$$

we approximate (2) by defining the sum

$$S_0 = \sum_{m=1}^n s(m) \frac{n}{n-m} e^{-b(\sqrt{n}-\sqrt{n-m})} \quad (5)$$

To evaluate S_0 , use the fact that

$$\frac{m}{2\sqrt{n}} < \sqrt{n} - \sqrt{n-m} < \frac{m}{2\sqrt{n-m}}, \quad 1 \leq m \leq n$$

and define the associated sums,

$$S_1 = \sum_{m=1}^n s(m) e^{-\frac{bm}{2\sqrt{n-m}}} \quad \text{and} \quad S_2 = \sum_{m=1}^n s(m) \frac{n}{n-m} e^{-\frac{bm}{2\sqrt{n}}}$$

whence $S_1 < S_0 < S_2$.

As $n \rightarrow \infty$, S_1 and S_2 (and so S_0) both converge to the same limit⁹, viz. the limit of

$$\sum_{m=1}^n s(m) a^m \quad \text{where} \quad a = \exp\left(-\frac{b}{2\sqrt{n}}\right) \quad (6)$$

It remains to investigate the behaviour of the series (6) as $n \rightarrow \infty$. To do this, it is convenient to write (6) in its alternative form (analogous to first summing within rows in Table 5), viz.

$$\sum_{m=1}^n m^2 \left[\sum_{k=1}^{\lfloor n/m \rfloor} a^{mk} \right] \quad (6')$$

The inner sum in (6') can be written as

$$a^m \frac{1 - (a^m)^{\lfloor n/m \rfloor}}{1 - a^m} \approx a^m \frac{1 - a^n}{1 - a^m}$$

whence (6') becomes

$$(1 - a^n) \sum_{m=1}^n \frac{m^2 a^m}{1 - a^m} \quad (7)$$

⁹ To show this: choose any θ , $0 < \theta < 1$ and break each of the sums S_1 , S_2 , S_0 into two parts,

viz. $\sum_{m=1}^{\lfloor \theta n \rfloor} (\cdot) + \sum_{m=\lfloor \theta n \rfloor + 1}^n (\cdot) = \tilde{S}_i + R_i$. The presence of the exponential term ensures that the

remainder term $R_i \rightarrow 0$ with $n \rightarrow \infty$. Moreover, by choosing θ sufficiently small, the ratio \tilde{S}_1/\tilde{S}_2 can be made arbitrarily close to 1.

For large n , we can approximate the sum in (7) by the integral

$$\int_1^n z^2 \frac{a^z}{1-a^z} dz = \int_1^n z^2 \frac{e^{-cz}}{1-e^{-cz}} dz$$

where $c = -\ell na = b/(2\sqrt{n})$. On writing cz as x this becomes

$$\frac{1}{c^3} \int_c^{cn} x^2 \cdot \frac{e^{-x}}{1-e^{-x}} dx = \left(\frac{2\sqrt{n}}{b} \right)^3 \int_{b/2\sqrt{n}}^{b\sqrt{n}/2} x^2 \frac{e^{-x}}{1-e^{-x}} dx$$

so that (7) becomes

$$\left(1 - e^{-b/2\sqrt{n}} \right) \left(\frac{2\sqrt{n}}{b} \right)^3 \int_{b/2\sqrt{n}}^{b\sqrt{n}/2} x^2 \frac{e^{-x}}{1-e^{-x}} dx \quad (8)$$

For large n , the integral in (8) converges to the standard integral

$$\int_0^{\infty} x^2 \frac{e^{-x}}{1-e^{-x}} dx$$

whose value is the constant

$$2 \sum_{k=1}^{\infty} \frac{1}{k^3} \approx 0.404$$

Writing this constant as c_0 , (8) coincides asymptotically with

$$c_0 \left(\frac{2\sqrt{n}}{b} \right)^3 = c_0 \frac{3}{\pi^3} \sqrt{\frac{3}{2}} n^{1.5}$$

	value of mk									
	1	2	3	4	5	6	7	8	9	10
m=1	1	1	1	1	1	1	1	1	1	1
m=2		4		4		4		4		4
m=3			9			9			9	
m=4				16				16		
m=5					25					25
m=6						36				
m=7							49			
m=8								64		
m=9									81	
m=10										100

Table 5. Table of coefficients of $p(n-mk)/p(n)$ in the sums (1) and (2): the table shows the case $n = 10$. The inner sum (over k) corresponds to summing a row in this table.

REFERENCES

Amaral, L.A.N., Buldyrev, S.V., Havlin, S., Salinger, M.A., and Stanley, H.E., (1998), 'Power law scaling for a system of interacting units with complex internal structure,' Physical Review Letters, vol. 80, pp. 1385-1388.

Andrews, G., (1970), The Theory of Partitions, Encyclopaedia of Mathematics and its Applications, vol. 2, Addison-Wesley, Reading, MA.

Fine, N.J., (1980), Basic Hypergeometric Series and Applications, Mathematical Surveys and Monographs no. 27, American Mathematical Society, Providence, Rhode Island.

Gibrat, Robert, (1931), Les inégalités économiques, applications: aux inégalités des richesses, à la concentration des entreprises, aux populations des villes, aux statistiques des familles, etc., d'une loi nouvelle, la loi de l'effet proportionnel, Paris: Librairie du Recueil Sirey.

Lee, Youngki, Luis A. Nunes Amaral, David Canning, Martin Meyer and H. Eugene Stanley, (1998), 'Universal Features in the Growth Dynamics of Complex Organisations,' Physical Review Letters, vol. 81, no. 15, 12 October, pp 3275-3277.

Pierou, V., Amaral, L.A.N., Gopikrishnan, P., Meyer, M. and Stanley, H.E., (1999), 'Similarities between the growth dynamics of university research and competitive economic activities,' Nature, vol. 400, pp. 433-437.

Stanley, Michael R., Luis A. Nunes Amaral, Sergey V. Buldyrev, Shlomo Harlin, Heiko Leschorn, Phillipp Maass, Michael A. Salinger and H. Eugene Stanley, (1996), 'Scaling Behaviour in the Growth of Companies,' Nature, vol. 319, 29 February, pp. 804-806.

Sutton, J, (1997), 'Gibrat's Legacy,' Journal of Economic Literature, vol. 35(1), March, pp. 40-59.