Edgar H. Sibley
Panel Editor

Taking Ein-Dor's recent reevaluation of Grosch's law one step further, the authors find evidence of different slopes for different classes of computers and the utility of an additional variable: the IBM or IBM-compatible factor. The analysis indicates that Grosch's law no longer applies to minicomputers.

COMMENTS ON "GROSCH'S LAW RE-REVISITED: CPU POWER AND THE COST OF COMPUTATION"

YOUNG MOO KANG, ROBERT B. MILLER, and ROGER ALAN PICK

As hardware costs continue to decline and more attention is paid to the economics of software development and operation, hardware nonetheless remains a significant expense. In the face of decisions to be made about centralization versus decentralization of CPU resources, knowledge about economies of scale can help organizations narrow down the choice of alternatives to be seriously considered. For this reason, Ein-Dor's recent reevaluation of the applicability of Grosch's law [1] sparked our interest. Having looked into his analysis, we believe there is yet more information contained in the data Ein-Dor uses than his analysis reveals. Specifically, we find evidence of different slopes for different classes of computers, and the utility of an additional variable—the IBM or IBM-compatible factor. We hope the interest generated by the original article [2, 5] will maintain itself for this further analysis of the same data.

GROSCH'S LAW

Grosch's law is a well-known statement about economies of scale in computing. An early expression of it by Grosch [3] is "I believe that there is a fundamental rule . . . giving added economy only as the square root of the increase in speed—that is, to do a calculation ten times as cheaply you must do it one hundred times as fast." In other words, Grosch's law is an assertion that the cost per MIPS, p, is related to power, w, by an affine function, f, as in $p = f(w^{-0.5})$. From an economic viewpoint, the statement is surprising since it implies increasing marginal productivity when decreasing marginal productivity is the more typical case with most technologies. The truth of this statement has been established by studies over the years and is often used as a justification for large centralized computing.

Motivated by the increasing prevalence of decentralized computing, often involving microcomputers, Ein-Dor [1] reexamined recent data to see if Grosch's law still applies. First, he regressed average cost against an intercept and computer power to produce a Grosch coefficient estimate of 0.30, which contradicts Grosch's law. Based on an observation that different sizes of computers may constitute nonhomogenous products and different technologies, he next separated computers into five categories: microcomputers, minicomputers, small mainframes, large mainframes, and supercomputers. Ein-Dor then performed a regression analysis to estimate the Grosch coefficient (exponent) with different intercept terms

© 1986 ACM 0001-0782/86/0800-0779 75¢

for each category. This resulted in a Grosch coefficient estimate of −0.55, which is close to Grosch's assertion that the value is −0.50. Ein-Dor concluded that, when computers are grouped according to their size and power, Grosch's law holds *within* the categories, but not *between* categories; he did not, however, pursue the question of the size of economies of scale within each individual category.

There is some controversy about the relevance of Grosch's law to the ongoing centralization/decentralization debate. Citation [5] is typical of much of this debate, while many more opinions have also appeared in the trade journals. For a scholarly perspective on this debate, see King [6]. Although the economics of the CPU represents only one part of the total debate, we nonetheless assert its importance. Large organizations continue to spend substantial sums upon CPUs.

IBM COMPATIBILITY

Is IBM so significant a force on the market for hardware that it constitutes a separate and different market from that of all other vendors? The distinction between IBM or IBM-compatible and non-IBM equipment was drawn in Ein-Dor's original data source [4], but not discussed in his article. In examining this question, we found that using a dummy variable to distinguish IBM machines from non-IBM machines did not result in a coefficient significantly different from zero. On the other hand, the use of a dummy to distinguish IBM and IBM-compatible machines from non-IBM-compatible machines results in a significant difference: All else being equal, an IBM or IBM-compatible CPU will cost more. We find that pooling IBM and IBM compatibles makes sense. In fact, Henkel [4] has argued that "It has become difficult to separate the PCM makers from IBM. In many ways, the PCM makers could be viewed more as extensions of IBM than IBM competitors."

TECHNOLOGICAL CHANGE

We suspected that Ein-Dor's analysis might be flawed by the fact that his data set—although taken from one point in time—included machines introduced in several different years and therefore represented several different technologies. To test this supposition, we accumulated introduction dates for most of the machines in the data set from a variety of sources and tried to determine the effect of technological change upon computer price: We found no statistically significant effect. Although we still believe technology to be an important factor in computer pricing, its significance is not revealed by straightforward statistical analysis—and we are continuing to study the question.

ON UNPOOLED GROSCH COEFFICIENTS

In experimenting with Ein-Dor's data, we found that, in addition to a set of distinct intercepts for the model, it is also appropriate to consider a set of distinct slopes. Ein-Dor found that assigning computers to five different categories according to price would produce different intercept terms. We extended this categorization to determine whether or not different Grosch coefficients would be found for each category. This would also give more flexibility by allowing different slopes for each category. (Ein-Dor's analysis had assumed the same slope for each category.)

Our analysis shows that the coefficients Ein-Dor presented were indeed improperly pooled. When an F-test was conducted, we found that there are distinct slopes within each category. The test procedure we followed is outlined on pages 94–95 of Neter, Wasserman, and Kutner [7]. The full model is displayed below in (3), while the reduced model is (3) with $b_1 = b_2 = b_3 = b_4 = b_5$. The test statistic

$$F^* = \{[SSE(R) - SSE(F)] \div [df(R) - df(F)]\} \div \{SSE(F) \div df(F)\} = \frac{9.703 - 8.216}{99 - 95} \div \frac{8.216}{95} = 4.3.$$

Comparing this statistic with the tabled F values $F_{.01}(4, 60) = 3.65$ and $F_{.01}(4, 120) = 3.48$ leads us to reject the hypothesis that all the stated coefficients are equal. For a detailed discussion, see [7, pp. 278–282].

THE MODEL

Ein-Dor's original model is given by

$$\log p = a + b \times \log w + k_1 d_1 + k_2 d_2 + k_3 d_3 + k_4 d_4. \tag{1}$$

A slight change in notation caused by adding a dummy term and removing the constant term results in the clearer but equivalent

$$\log p = b \times \log w + k_1 d_1 + k_2 d_2 + k_3 d_3 + k_4 d_4 + k_5 d_5. \tag{2}$$

Our model is given by

$$\log p = a \times i + b_1 \log w_1 + b_2 \log w_2 + b_3 \log w_3 + b_4 \log w_4 + b_5 \log w_5 + k_1 d_1 + k_2 d_2 + k_3 d_3 + k_4 d_4 + k_5 d_5 \tag{3}$$

TABLE I. Estimation Results

Estimation results with dummy for IBM compatibles and different slopes for a different class of computers. Adjusted $R^2 = 1 - (1 - R^2) \times (N - 1)/df$, where $N = 106$ and $df = 95$, 0.937				
Variable	Coefficient	Standard error	T	P(2 − tail)
i	0.152	0.066	2.31	0.023
$d_1 \log w$	−0.694	0.286	−2.43	0.017
$d_2 \log w$	−0.433	0.089	−4.88	0.000
$d_3 \log w$	−0.387	0.087	−4.47	0.000
$d_4 \log w$	−0.901	0.108	−8.36	0.000
$d_5 \log w$	−0.702	0.159	−4.42	0.000
d_1	7.581	0.718	10.55	0.000
d_2	6.517	0.124	52.41	0.000
d_3	5.689	0.059	95.84	0.000
d_4	4.240	0.123	34.57	0.000
d_5	1.656	0.195	8.49	0.000

Analysis of variance					
Source	Sum of squares	df	Mean square	F-ratio	P
Regression	3286.686	11	298.790	3454.862	0.000
Residual	8.216	95	0.086	—	—

where

$$\log w_j = d_j \times \log w, \quad j = 1, 2, 3, 4, 5 \quad \text{and}$$

$$i = \begin{cases} 1, & \text{if the machine is IBM compatible} \\ 0, & \text{otherwise} \end{cases}$$

$$d_1 = \begin{cases} 1, & \text{if machine is a supercomputer} \\ 0, & \text{otherwise} \end{cases}$$

$$d_2 = \begin{cases} 1, & \text{if machine is a large mainframe} \\ 0, & \text{otherwise} \end{cases}$$

$$d_3 = \begin{cases} 1, & \text{if machine is a small mainframe} \\ 0, & \text{otherwise} \end{cases}$$

$$d_4 = \begin{cases} 1, & \text{if machine is a minicomputer} \\ 0, & \text{otherwise} \end{cases}$$

$$d_5 = \begin{cases} 1, & \text{if machine is a microcomputer} \\ 0, & \text{otherwise} \end{cases}$$

The values of the parameters b_1, b_2, b_3, b_4, and b_5 are the Grosch coefficient for each class of computer.

Estimation results are given in Table I; after incorporating these results, the model becomes

$$\begin{aligned}\log p = {} & 0.152i - 0.694 \log w_1 \\ & - 0.433 \log w_2 - 0.387 \log w_3 \\ & - 0.901 \log w_4 - 0.702 \log w_5 + 7.581 d_1 \\ & + 6.517 d_2 + 5.689 d_3 \\ & + 4.240 d_4 + 1.656 d_5. \end{aligned} \quad (4)$$

CONCLUSIONS

Our further analysis largely supports Ein-Dor's conclusions, giving added strength, we believe, to the continuing validity of economies of scale in CPU power. It indicates that there is still reason for organizations to consider centralized CPUs when other factors are equal: more specifically, that there is some economic force continuing to favor the use of large CPUs rather than many small CPUs of equivalent aggregate power. We feel that for large data sets the separation of slope coefficients among classes of computers should not be ignored. Finally, we wish to point out that the consideration of varying slopes results in a Grosch coefficient closer to one than to one-half for computers in the minicomputer class, meaning that, at least for one class of computer, Grosch's law is no longer valid.

REFERENCES

1. Ein-Dor, P. Grosch's law re-revisited: CPU power and the cost of computation. *Commun. ACM 28*, 2 (Feb. 1985), 142–151. The most recent scholarly study of Grosch's law.
2. Ein-Dor, P. Rejoinder to response on re-revisitation (of Grosch's law). *Commun. ACM 28*, 7 (July 1985), 668. A response to [5].
3. Grosch, H.R.J. High speed arithmetic: The digital computer as a research tool. *J. Opt. Soc. Am. 43*, 4 (Apr. 1953), 306–310. An early statement of Grosch's law.
4. Henkel, T. Hardware roundup: A look at 98 systems from 17 vendors. *Computerworld 15*, 26 (July 13, 1981), 12. Ein-Dor's data source.
5. Jones, J. On Grosch's law revisited. *Commun. ACM 28*, 5 (May 1985), 453. Argues that Grosch's law is less important than it once was since it ignores behavioral considerations.
6. King, J.L. Centralized versus decentralized computing: Organizational considerations and management options. *Comput. Surv. 15*, 4 (Dec. 1983), 319–349. An authoritative survey that puts Grosch's law into perspective with respect to the considerations relevant to the larger centralization versus decentralization debate.
7. Neter, J., Wasserman, W., and Kutner, M. *Applied Linear Statistical Models*. Richard D. Irwin, Homewood, Ill., 1985. A standard textbook.

CR Categories and Subject Descriptors: C.4 [**Performance of Systems**]; K.1 [**The Computing Industry**]; K.6.0 [**Management of Computing and Information Systems**]: General

General Terms: Economics, Performance

Additional Key Words and Phrases: decentralization, economies of scale, Grosch's law, mainframes, microcomputers, minicomputers, supercomputers

Received 12/85; accepted 2/86

Authors' Present Address: Young Moo Kang, Robert B. Miller, and Roger Alan Pick, School of Business, University of Wisconsin-Madison, 1155 Observatory Drive, Madison WI 53706. As of Sept. 1, Roger Alan Pick's address will be as follows: Quantitative Analysis Systems and Information Dept., Lindner Hall (ML 130), University of Cincinnati, Cincinnati, OH 45221.

Permission to copy without fee all or part of this material is granted provided that the copies are not made or distributed for direct commercial advantage, the ACM copyright notice and the title of the publication and its date appear, and notice is given that copying is by permission of the Association for Computing Machinery. To copy otherwise, or to republish, requires a fee and/or specific permission.