1-1-2003

A Rule Based Forecast of Hard Disk Drive Costs

G. Kent Webb
San Jose State University, g.webb@sjsu.edu

A RULE BASED FORECAST OF COMPUTER HARD DRIVE COSTS

G. Kent Webb, San Jose State University, webb_k@cob.sjsu.edu

ABSTRACT

While the cost of storage devices such as hard disk drives continues to fall, the overall proportion of computer network costs dedicated to storage continues to rise. Within a few years, storage will account for 50 percent of total network hardware and software costs. Since computer networks typically have long lives, the design process often involves projecting costs many years into the future. This paper examines a rule based on a technology trend that can be used to estimate the cost per megabyte of hard disk drives. The rule is similar to the well-known Moore's law that has reliably summarized integrated circuit advances over the past several decades. A statistical analysis of historic data suggests the rule for hard disk drives captures much of the readily available information.

Keywords: Rule Based Forecast, Cost Estimation, Technology Trends, Hard Disk Drive

INTRODUCTION

Perhaps the most famous rule for forecasting the pace of technical development is Moore's law. When proposed in 1965 (9) Moore was the Director of Research Labs at Fairchild Semiconductor and co-founded Intel three years later. He observed an exponential growth path on the number of transistors per integrated circuit and predicted that this trend would continue, doubling about every 18 months. Although exponential growth paths cannot continue indefinitely, Moore's law provided a remarkably reliable forecast for nearly the past four decades and is expected to prevail for at least another decade (10).

A similar rule summarizing the pace of hard disk drive development was observed by engineers at IBM's Disk Drive Research Center in San Jose, California. The rule suggests that storage densities will increase by 60 to 100 percent per year over the next 25 years with a resulting 40 to 45 percent annual decline in the price per megabyte of disk storage (5), a somewhat faster pace of advancement than proposed by Moore's law for integrated circuits. Given the increasing importance of storage in computer networks, this rule based cost estimate can be useful in network analysis and design.

To provide some background for the problem, the next section of the paper reviews some of the literature on rule based forecasting. Following that is a section presenting the technical basis for the proposed rule. A statistical analysis of historic data from 1980 to 2002 is then presented to calibrate the rule based forecast and explore some market indicators. Conclusions related to limitations, future research, and other related technologies complete the paper.

Rule Based Forecasting

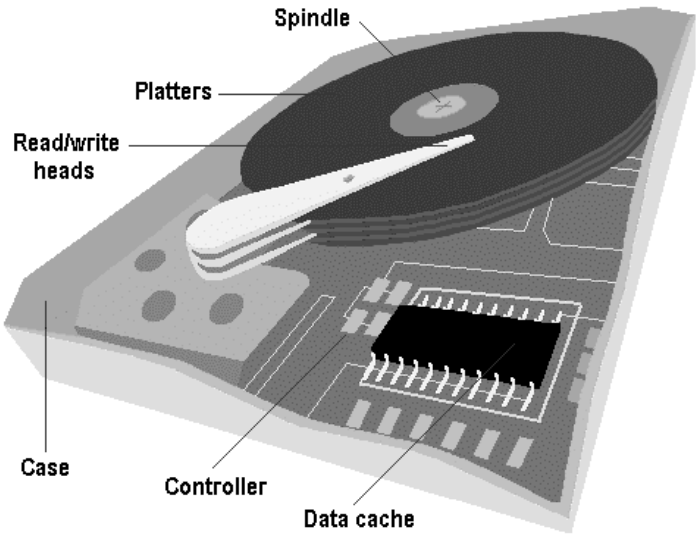
There are a variety of approaches to rule based forecasting described in the literature. Armstrong, with colleagues, proposes that: "rules are used to combine forecasts from simple extrapolation methods" (1). Also, that it involves the development of an "expert system that integrates judgment and statistical procedures to combine forecast" (2). Clarkson, entrenched in a telecommunications technology background, defines rule-based systems as including artificial intelligence, fuzzy logic, and genetic algorithms (3).

A number of studies attempt to present general rules for technology forecasting, similar to the one described in this paper. One researcher proposes a rule that hardware prices decrease at the rate of about 20 to 25 percent per year, while software costs seem to follow no rules (11). Another study shows that electronic component prices fall at about 15 percent per year over the first few years of the product life, and at a lower rate thereafter (14). A recent general forecast of computer technology examines where exponential rules such as Moore's law can be applied (4). Economists have developed a technique called cointegration to evaluate the long-run stability of these kinds of rules. (7, 8, 12)

Hard Disk Drive Technology

As illustrated in Figure 1, hard disks are rigid platters, composed of a substrate and a magnetic medium. The substrate - the platter's base material – must be non-magnetic and capable of being machined to a smooth finish. It is made either of aluminum alloy or a mixture of glass and ceramic. To allow data storage, both sides of each platter are coated with a magnetic medium - formerly magnetic oxide, but now, almost exclusively, a layer of metal called a thin-film medium. There's a gap between the platters, making room for magnetic read/write head, mounted on the end of an actuator arm.

Figure 1: Inside the Hard Disk Drive



As illustrated in figure 2, two or more platters are stacked on top of each other with a common spindle that turns the whole assembly at several thousand revolutions per minute. Data is recorded onto the magnetic surface of the disk in exactly the same way as it is on floppies or digital tapes. Essentially, the surface is treated as an array of dot positions, with each "domain' of magnetic polarization being set to a binary "1" or "0".

Storage capacity depends on the size of the platters and the density of storage. Usually measured in bits per square inch (bpsi), referred to as the areal density, the amount of data that is stored on a disk per square inch is equal to the tracks per inch multiplied by the bits per inch along each track. Increasing this areal density facilitates the rule that prices will fall at about 45 percent per year over the next several years. Magnetoresistive (MR) Head Technology, invented by IBM, provides the technical foundation for current drive advances. Table 1 below summarizes some of the relevant developments related to disk drive technology.

Figure 2: Stacked Platters Inside the Hard Disk Drive

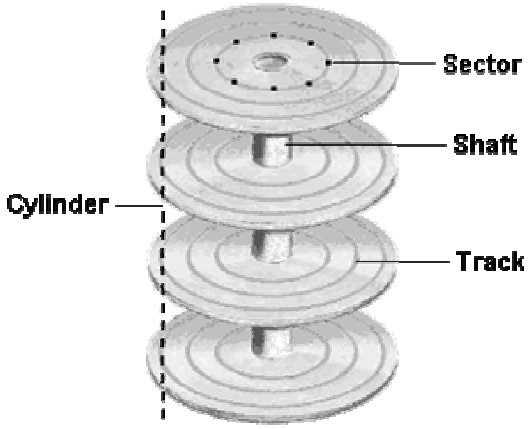


Table 1: Milestones in the Development of Disk Drives

1954	IBM invents the hard disk drive with a storage capacity of 5 Megabytes (MB) stored across fifty 24 inch platters
1956	The drive is commercially available for a monthly rental of $3,200. Purchase price is $50,000, $10,000 per megabyte in current dollars.
1963	Platter size reduced to 14 inches
1973	Winchester Drive Introduced (named after the project manager's gun), the first closed environment drive (still used)
1979	Al Shugart forms a company to make 5.25 inch hard drives for the personal computer
1987	77 Hard drive manufacturers (peak)
1989	Storage Density Reaches 1 Gigabyte per square inch (a gigabyte will hold about 80,000 printed pages)
1995	Storage Density Reaches 3 Gigabytes per square inch
2001	Glass substrates developed for the platters increases areal density to 10 Gigabits per square inch
2002	16 Hard drive manufacturers. Storage cost is about $.0029 per megabyte, less than 1/3 of a cent.
2003	Projected 400 Gigabyte drive for PC

DATA ANALYSIS

The Data

Prices and capacities of newly introduced disk drives are maintained on an independent website (6). Data on orders and shipments of computer equipment are collected by the Department of Commerce (13). Data span the period from 1979 to 2002 and represent annual averages.

The Results

Two regression equations are summarized in the table 2. Equation 1 correlates the dependent variable, percent change in price per megabyte for disk drives, to an average (the intercept), market conditions summarized by the percentage change in new orders for computer equipment, and a dummy variable to separate the data into pre-1990 and 1990 and later. The t-statistic reported for each variable indicates that the intercept and dummy variable are highly significant, but that the impact of new orders is tenuous. The positive sign of the coefficient for the new orders variables conforms to logic, in expanding markets technology advances faster. Given the low statistical correlation between percentage change in new orders and percentage change in drive price per megabyte, another equation was estimated using only the time trend elements.

Equation 2 suggests that the average percentage change in the cost per megabyte for disk drives from 1980 to 1989 was 18.675 percent. From 1990 forward, the percentage change was 48.254 percent.

The null hypothesis that there is no correlation between percentage change in the price per megabyte of drives and the percentage change in new orders for computer equipment is accepted. The null hypothesis that there is no correlation between the time trend from 1980 to 1989, and a different time trend from 1990 to 2002 is rejected.

Table 2: Regression Analysis of Historic Data

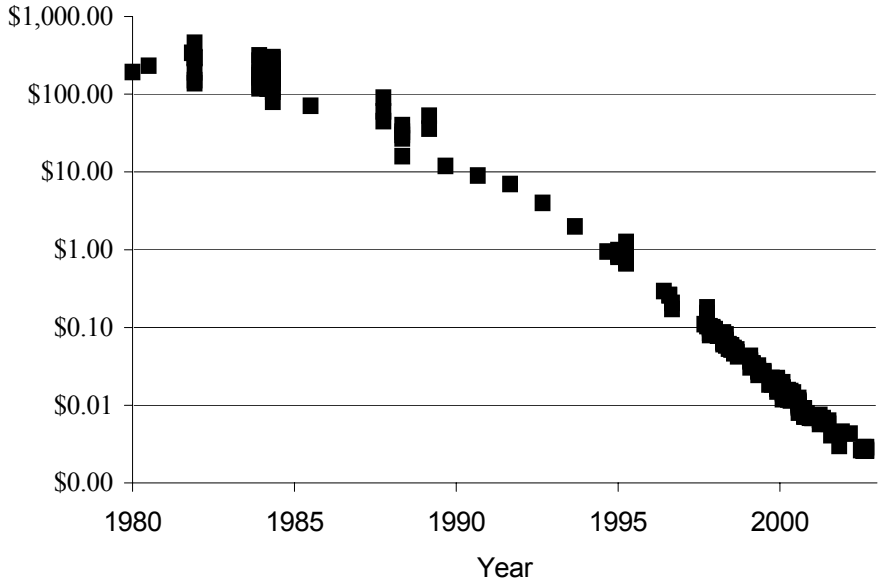
Dependent Variable: Percent Change in Price Per Megabyte for Hard Disk Drives

Equation 1: Independent Variables (Annual Data)	Coefficient	t-statistic	P-value
Intercept	-0.17646	-3.13736	0.005424
Percentage Change in Orders for Computer Equipment	-0.10183	-0.40627	.689084
Dummy Variable Equal to One for 1990 and later	-0.29941	-4.83854	0.000114
R-Square = 0.555			
Equation 2: Independent Variable (Annual Data)	Coefficient	t-statistic	P-value
Intercept	-0.18675	-3.7982	0.001127
Dummy Variable Equal to One for 1990 and later	-0.29479	-4.95082	0.00005
R-Square = 0.551			

Graphical Results

Figures 3 and 4 show the logic of the regression analysis. Figure 3 shows the change in the price per megabyte on a logarithmic scale. The nearly straight line supports the exponential pattern of declining prices for disk storage. Figure 4 shows the percentage change in new orders for computer equipment compared to the percentage change in price per megabyte. Little correlation is apparent in the graph, as is indicated in equation 1 in the table above.

Figure 3: Log Scale for Price Per Megabyte for New Hard Drives 1980 to 2002



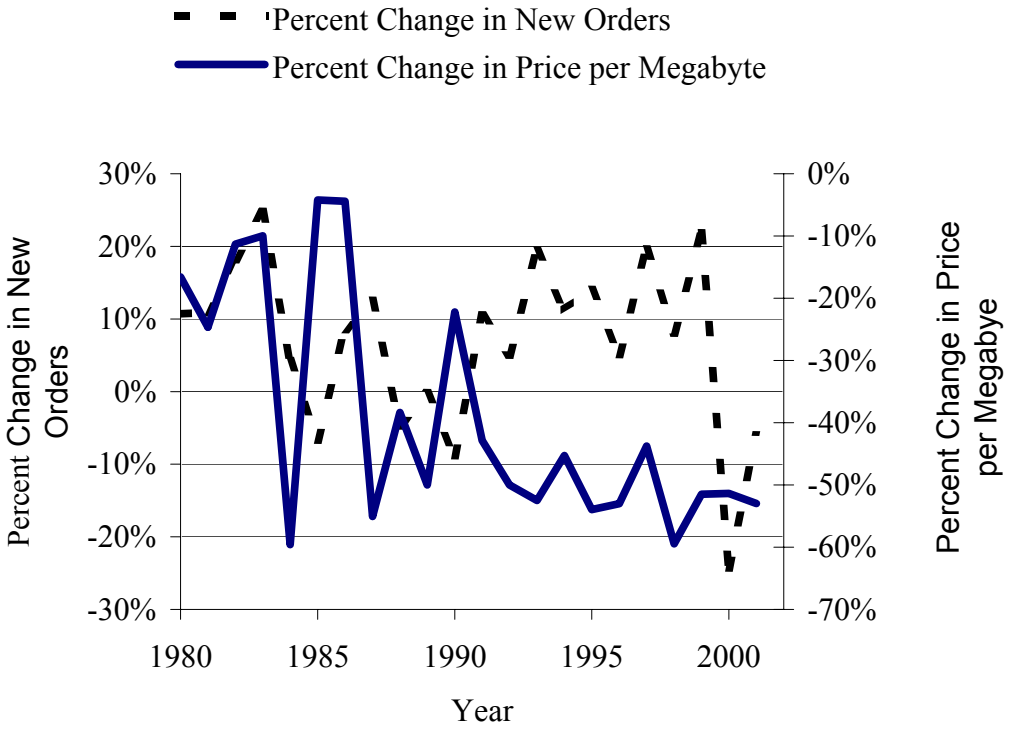
CONCLUSION

A useful technical rule for forecasting future hard drive costs is supported by the data and analysis presented in this paper. The data suggest that if present trends continue, the cost per megabyte of hard disk drive storage will fall at a rate of about 48 percent per year. This rate is slightly faster than the 40 to 45 percent rate suggested by IBM engineers who have observed an approximate annual doubling of areal storage density. The engineers felt that this trend could continue for 25 years or more. Theoretical limits allow for a substantially longer advance in storage density.

Many developing storage technologies, such as RAID (redundant array of inexpensive disks) and SAN (storage area network) are based on hard disk drives. The percentage rate of decline in cost for disk drives can therefore be approximately applied to these developing technologies.

The impact of short-term changes in market conditions, specifically the percentage rate of change in new orders for computer equipment, proved to be statistically insignificant, although there was a hint of correlation. Given the difficulty of forecasting future orders of computer equipment, this approach provides little additional help in making cost forecasts. Future research might consider other market variables such as market size, the number of firms, and the number of new product introductions. Other emerging technologies such as optical storage might be interesting candidates for analysis.

Figure 4



REFERENCES

1. Adyaa, Monica; J. Scott Armstrong, Fred Collepy, Miles Kennedy (2000). An Application of Rule-Based Forecasting To a Situation Lacking Domain Knowledge, International Journal of Forecasting, (16) 477-484.
2. Adyaa, Monica; J. Scott Armstrong, Fred Collepy, Miles Kennedy (2001). Automatic Identifcation of Time Series Features for Rule Based Forecasting, International Journal of Forecasting, (17), 143-157
3. Clarkson, Tevor (1999). Applications of Neural Networks in Telecommunications, Working paper, King's College (available at: http://www.erudit.de/erudit/events/tc-d/16_04_99/1_clarkson.pdf)

4. Gray, Jim (2000). Computer Technology Forecast for Virtual Observatories: Extended Abstract of talk at Astronomy Virtual Observatories of the Future, California Institute of Technology (available at: http://research.microsoft.com/~Gray).
5. Grochowski, Ed (1994). The Era of Magnetoresistive Heads, IBM Research Paper. (Available at: http://www.storage.ibm.com/ipl/oem/tech/eraheads.pdf)
6. Historical Notes about the Cost of Hard Drive Storage Space (available at: http://www.alts.net/ns1625/winchest.html).
7. Johansen, S and K. Julselius (1990). Maximum Likelihood Estimation and Inference Cointegration – With Applications to the Demand for Money, Oxford Bulleting of Economics and Statistics (52), 169-210.
8. Johansen, S. and K. Juselius (1992). Testing Structural Hypothesis in a Multivariate Cointegration Analysis of the PPP and the UIP for UK, Journal of Econometrics (53) 211-422.
9. Moore, Gordon (1965). Cramming More Components Onto Integrated Circuits, Electronics, (38) 8. (Available at: http://www.intel.com/research/silicon/moorespaper.pdf)
10. Moore, Gordon (2003). No Exponential Is Forever…But We Can Delay 'Forever'" International Solid State Circuits Conference. (The slideshows from this presentation are available on Intel's Website at: ftp://download.intel.com/research/silicon/Gordon_Moore_ISSCC_021003.pdf).
11. Nilles, Jack M. (1998). Managing Telework: Strategies for Managing the Virtual Workforce. John Wiley & Sons, Inc.
12. Riis, Jacob (1996). Forecasting Danish Timber Prices with an Error Correction Model, Journal of Forecast Economics (2)3, 257-271.
13. U.S. Cenus Bureau, Manufacturing, Mining, and Construction Statistics (available at: http://www.census.gov/indicator/www/m3/hist/naicshist.htm).
14. Webb, G. Kent (1994). Integrated Circuit Pricing, the Journal of High Technology Management Research (5) 247-260.