RAPID CHANGE IN THE PERSONAL COMPUTER MARKET

A QUALITY-ADJUSTED HEDONIC PRICE INDEX

1976 - 1987

by

JEREMY MICHAEL COHEN

B.S., Computer Science
Brown University
(1983)

M.S., Computer Science
Union College
(1986)

Submitted to the Sloan School of Management
in Partial Fulfillment of
the Requirements of the Degree of
Master of Science in Management

at the

Massachusetts Institute of Technology

May 1988

Signature of Author____________________________
Alfred P. Sloan School of Management
May 13, 1988

Certified by____________________________
Ernst R. Berndt
Professor of Applied Economics
Thesis Supervisor

Accepted by____________________________
Jeffrey A. Barks
Associate Dean, Master's and Bachelor's Programs

RAPID CHANGE IN THE PERSONAL COMPUTER MARKET

A QUALITY-ADJUSTED HEDONIC PRICE INDEX

1976 - 1987

by

Jeremy Michael Cohen

Submitted to the Alfred P. Sloan School of Management
on May 13, 1988 in partial fulfillment
of the requirements for the Degree of
Master of Science in Management

ABSTRACT

Many economists feel that the government consumer and producer price indexes may be overestimating the effect of inflation since the indexes likely do not account properly for the extremely rapid technological advances in many areas.

One of these areas is personal desktop computers. Although personal computers have only been available for approximately ten years, their technical quality characteristics (RAM, clock speed, disk storage, etc.) have improved by several orders of magnitude. Their prices, however, while dropping significantly, have not dropped nearly as fast as their quality characteristics have increased.

Thus, a method is needed to isolate the pricing effects from the quality change effects. One way to do this is to create an Hedonic Price Index. The hedonic hypothesis states that a computer transaction is a tied sale of a bundle of quality characteristics, with the price of each computer model directly related to the quantities of characteristics contained within it.

In this thesis I compute a quality-adjusted hedonic price index for personal computers (PCs), using over 1100 time series and cross-sectional data points with PC models from over 110 manufacturers. In addition to computing an overall price index, I examine the different pricing policies of several PC manufacturers (IBM, Apple, Compaq, NEC, etc.).

The data set is also stratified in several ways to compute price indexes for different subsets of PCs. Further, the value each quality component adds to the overall PC system is computed to determine the relative effect of each component on the overall PC system prices.

The personal computer market has been very dynamic. In part this is shown in the final result of an annual average quality-adjusted real price decline of 25.3%. This corresponds to a PC that would have cost over $26,000 in 1976 costing only about $1,000 in 1987. In addition, the rate of decline has varied considerably and this trend of volatility is analyzed to see what implications it may have for PC consumers and manufacturers in the future.

Thesis Supervisor: Dr. Ernst R. Berndt
Title: Professor of Applied Economics

Table of Contents

Acknowledgements

I would like to thank my thesis advisor Prof. Berndt and my thesis reader Prof. O'Brien for their help and guidance as I worked on this thesis. The time and effort they have put in as well as their enthusiasm and insights helped immeasurably. I would also like to thank Prof. Berndt for teaching such a fascinating Econometrics class that set me on the road leading to this thesis.

I would also like to thank my parents for their steadfast support over the years. I give them much of the credit for what I have done so far and what I will do in the future.

Finally, I wish to thank my fiancée, Penney Bigel, both for her assistance in creating this thesis and for her warmth, her smile, and for saying "yes"...

1.0 Introduction

In 1986 the Bureau of Economic Analysis within the Department of Commerce began using hedonic price indexes of computers for deflation within the National Income and Product Accounts. This explicitly stated the government's belief that the prices of computer systems had indeed been decreasing since their introduction in 1953.

Prior to 1986 it was assumed that real computer prices had been completely constant; i.e., that the rate of price change was exactly equal to that of the Consumer Price Index. This conclusion was contradicted, however, both by detailed economic analysis and by market examination. The BEA change was made to account better for the extremely rapid technological advances in many areas of the computer market.

These rapid advances have led to steadily declining cost/performance ratios for computer systems. For example, one recent study estimated that the cost in 1984 of an equivalent 1953 mainframe would be 1/10 of 1%, a dramatic cost/performance change.[1] Several studies have recently been published that examine this phenomenon within the high-end computer mainframe and peripheral markets; these are described further in Chapter 2.

One pundit nicely described the magnitude of the price decline in computer mainframes as follows:

> If the auto industry had done what the computer industry has done in the last 30 years, a Rolls-Royce would cost $2.50 and get 2,000,000 miles to the gallon.[2]

While the mainframe market has been studied extensively, no comprehensive study to date has examined the price/performance trend within low-end, personal

---

[1] Triplett, p. 1.

[2] Forbes, December 22, 1980, p. 24, attributed to Computerworld.

computers, a rapidly growing segment of the computer market.

Some empirical results have been obtained, such as one small study which found a quality-constant price decrease of PCs from 1982 to 1986 between 20 and 25%.[3] Another study briefly mentioned that the rate of price decline of PCs declined at the same steep rate as larger systems between 1972 and 1982 and declined even faster during 1983 and 1984.[4]

This effect was echoed by William Lowe, IBM Vice President and President of the Entry Systems Division, who stated at the March, 1988, Boston Computer Society General Meeting that IBM PC price declines have been running over 22% for the last few years. This price decline trend can also be seen in the "Computer Blue Book," which lists the original and current prices for a great variety of computers. One example of this PC price decline is that the price for a Lisa 1 computer originally sold in 1984 for $6,995 is now $626.[5]

Thus, it appears that the price/performance ratio for personal computers has been declining at least as fast as the computer mainframe ratio has been declining. This thesis examines this hypothesis using regression analysis of a large (1108 entries, 406 PC models, 114 manufacturers) data base containing entries from the start of the PC market in 1976 through 1987.

To examine the changing price/performance ratio, a method is needed to isolate the pricing effects from the quality change effects. One fairly simple technique involves creating a "matched model index". This index is formed from the change in prices for identical PC models in adjacent years. This holds the technology or performance constant and only examines the change in price for each model.

---

[3] Gordon, p. i.

[4] Cartwright, p. 9.

[5] Bell, p. 42.

While this is relatively easy to compute, the technique does have several problems. First, it works best when prices are stable and the number of new models is small. This is certainly not the case in the extremely dynamic PC market.

Second, newer models with better technology and greater performance having increased performance are completely ignored by the matched model index. This will often cause the index to underestimate the magnitude of the decline of the quality-adjusted prices.[6]

Third, if the models matched in the index are not identical the index results will be biased. This is definitely a problem in the PC market since many PC models are upgraded on a fairly frequent basis.

The matched model index basically assumes that the prices of models embodying older technology change instantaneously such that their quality-adjusted prices are equal to models with newer technology.[7] While this may well be true for stable products such as commodities, it is less appropriate for such rapidly improving products as PCs.

Thus, a technique is needed to account both for the explicit price declines of PC models, as well as the implicit price declines that occur as a result of new models coming out with better technology and greater performance.

One well-known procedure for dealing with quality change is the hedonic price technique. The hedonic hypothesis suggests that a computer transaction is a sale of a bundle of characteristics, with the price of each personal computer model reflecting the quantities and quality of the characteristics embodied in it. Thus, the hedonic technique relates the cost of the personal computer to the amount of its

---

[6] Dulberger, p. 49.

[7] Cole, p. 48.

quality characteristics or attributes. A succinct description of the hedonic price index method follows:

> Estimates of implicit... characteristics' prices are derived from estimates of characteristics' coefficients. These implicit characteristics' prices are then used to estimate the price of an unobserved model by valuing its embodied characteristics. In constructing a quality adjusted price index such estimates are then used for prices of models not transacted in the reference period.[8]

The hedonic method holds the quality-related characteristics constant and then examines the implicit price changes relating to increased performance, in addition to the effects of overall inflation. The index created is a mean of all the price changes by year, observed across the hedonic function.

One specific weakness of the hedonic model is that it holds constant only the most measurable and important characteristics in deriving the quality-adjusted price index.[9] I have addressed this issue in particular by attempting to have as thorough a set of quality variables as is possible, given the constraints of data collection.

In this thesis I compute a quality-adjusted (or quality-corrected) hedonic price index for personal computers. The PC market has been very dynamic with constant technological improvements, new model releases, and declining prices. It is expected, therefore, that the hedonic price index I derive will show very significant quality-adjusted price declines over the 1976 to 1987 time frame, at least as large as in the mainframe studies to be discussed in section 2.2.

The thesis is divided into three main sections. Following a survey of the literature, I review my data collection process and rationalize the choice of variables. Then, I describe the process used in creating the hedonic price index, and interpret

---

[8] **Op. cit.**, p. 1.

[9] Cartwright, p. 7.

the results. Finally, I conclude by examining some extensions of this work and the importance of the results.

2.0 Survey of Literature

Hedonic price studies have been used since the 1930's to analyze price level changes in various products. In this section, I first briefly discuss two early hedonic studies that were important in the development of the hedonic methodology. Second, I shall describe several of the recent hedonic price studies in the computer industry and examine what overall results they obtained.

2.1 Early Hedonic Studies

The first apparent empirical study relating quality and price using a hedonic technique was done by Frederick Waugh in the 1920's on the prices of asparagus, tomatoes and cucumbers.[10] He attempted to determine the relative values of certain quality characteristics (such as color, uniformity, size) of vegetables when factored into the overall price. While Waugh's study may only have limited applicability today, it is still an interesting example of the hedonic technique used for non-durable, non-technological products.

The first hedonic price study to examine price and quality changes over time was done by General Motors (GM) in the late 1930's.[11] GM had come under fire in Congress over whether it had been pricing its cars monopolistically. Apparent evidence cited by critics was a US Bureau of Labor Statistic's new car price index, indicating that prices of new GM cars had risen by 45% between 1925 and 1935.

The GM hedonic price study showed that while prices had risen by 45%, the quality-adjusted new car hedonic price index had decreased by 53%. Although GM had been raising the price of its new cars, the quality characteristics (weight,

---

[10] Frederick V. Waugh, "Quality Factors Influencing Vegetable Prices," Journal of Farm Economics, vol. 10, no. 2, April, 1928, pp. 185-196.

[11] "Hedonic Price Indexes with Automotive Examples," The Dynamics of Automobile Demand, General Motors Corporation, NY, 1939, pp. 99-117.

length, horsepower) of the new cars had risen far faster than the prices. This finding was quite dramatic and provided important input to the policy discussions. Thus, not only was this the first hedonic study to examine data over time but, perhaps more importantly, it set a precedent for hedonic price indexes to be used in general policy determination.

2.2 Hedonic Studies in the Computer Industry

In the last two decades several hedonic studies have been published, focusing on various segments within the computer industry including mainframes, peripherals, and minicomputers. In this section, I shall review several of these studies and examine the results they obtained.

Chow (1967) performed hedonic modelling on the rental price of minimally configured mainframes based on processor speed (multiplication time, memory access time) and main memory size. His results indicated that all three variables were quite significant, particularly memory size. He obtained an R^2 of .908 and an average annual growth rate (AAGR) of quality-adjusted nominal[12] prices of -20.8% for the 1960 to 1965 time frame.[13]

Cale (1979) studied the relationship between computer prices and hardware performance and tested whether Grosch's Law[14] was still applicable to the computer field. He examined general purpose and small business computers from 1970 to 1977 using memory (RAM) and Direct Access Storage Devices (DASD) as in-

---

[12] All of the price indexes listed in the Literature Survey chapter are nominal, as opposed to real. There is only little difference between the two scales, however, until the 1970s.

[13] R squared measures the proportion of the sample variance in prices "explained" by movements in the explanatory variables.

[14] Grosch's Law - "The power of a computer system increases as the square of its cost." Dr. Herbert Grosch. From Cole, p. 225.

dependent variables. The results indicated that both variables were quite significant (t-stats > 5) and the overall R^2 was .925. The results also demonstrated much more stability for the general purpose systems (R^2 = .854) than for the small business systems (R^2 = .626). The overall AAGR for the study was -8.5%.

Cole (1986) focused on mainframe computer components (processor and peripherals), as opposed to complete systems. The data used was from 1972 to 1984 and involved mainframe processors, disk drives, printers and displays. The variables used were speed, capacity, plus special technological variables meant to take into account various new technologies that came into use during this time frame. In addition, four price indexes were computed, including matched model (processor AAGR of -3%), composite hedonic (-17.8%), characteristic hedonic (-17.6%), and regression hedonic (-19.2%). Similar results were obtained for peripherals, though the AAGRs were in the -7% to -16% range for disk drives and printers and -1% to -8% for displays.

Another interesting observation by Cole was that the post-1977 AAGRs were much greater (more negative) for all components except disk drives, indicating that quality-adjusted prices had been dropping at an accelerating rate in recent years. This study formed the basis for the Department of Commerce's Bureau of Economic Analysis (BEA) official computer price index, first released in 1986.[15]

Cartwright (1986) evaluated the Cole study and the new BEA computer price index. Overall, he stated that the use of the new deflator for mainframe computers represented a major improvement over the previously used convention of assuming no price changes for computers or using an older matched model index.[16] He in-

---

[15] Berndt, chapter 4, p. 34.

[16] Cartwright, p. 9.

dicated, however, that problems exist with the new index due to limited sample size, use of list rather than transaction prices and incomplete shipping information.

Dulberger (1986) is a more in-depth report on the price indexes for computer processors first reported in the Cole survey. In addition to presenting a comparison of the four types of price indexes, Dulberger also examined the alternative functional form question and determined that the double log form appears best due to the lowest standard error of estimate for the Box-Cox transformation.

Gordon (1987) constructed new hedonic price indexes for computers from 1951 to 1984 using the Chow and Dulberger datasets plus a few others. The study examined both mainframes and super-mini computers and generated a price index with an AAGR of -19.8%. The study emphasized the weighting of computer price indexes to determine an overall index and illustrated how this index is substantially different from the BEA deflator during the 1957 to 1972 time period.

Gordon also computed a matched model price index using a very small PC data set. His results indicated that the 1981 to 1986 nominal price AAGR was -21.9%.

Triplett (1987) surveyed the quality-adjusted price indexes for computers from 1953 to 1984 and combined them to construct a "Best Practice" index. From 1953 to 1972 he obtained an AAGR of -21.2% and from 1972 to 1984 an AAGR of -13.1%. In addition, he showed how the yearly price declines are non-uniform and are quite well correlated with the advent of new computer generations (1964, 1972) and with the introduction of new technologies such as semiconductors. He also stated that capturing the full impact of technological change on computers requires the modelling of whole systems as opposed to working at the component level.[17]

[17] Triplett, p. 13.

As can be seen, extensive work has been done in the last several years in creating various hedonic price indexes for computer mainframe systems and components. To date, however, no full-fledged study has been published examining the relatively new personal computer market, although Gordon does give some very preliminary results.

3.0 Methodology

In this section I review the methodology of the research, including data collection and regression design.

3.1 Data Types

The two types of data I collected for this study were technical and pricing. The technical data consisted of the physical attributes ("quality" attributes) of each personal computer. This type of data was reasonably easy to obtain from magazine reviews and advertisements, but several difficulties emerged.

The first difficulty with the technical data collection was that it was much easier to find the technical data on the more recent (1985+) PC models than it was for the earlier models. In particular, prior to 1980 it was often difficult to discover some of the more esoteric technical attributes (such as the amount of ROM) of the PCs.

In cases involving very similar models, I initially thought of extrapolating some of the technical attributes from one model to another. Usually, however, I did not feel justified in doing this, so I generally left the appropriate data slot empty. This explains why, for example, I have only 325 entries for the ROM variable out of my 1108 total data points. As I will show later, however, this variable turned out not to be important.

Also, in a few cases there were near-identical versions of one model of a PC (such as the IBM AT models 319 and 339, which differed only by type of keyboard). In these cases I used only one of the models (the standard version), since using the other model also would not have added any new information to the regression but would have instead weighted it towards the duplicated model.

Overall, obtaining the technical data was not difficult. In comparison, however, obtaining the pricing data posed several serious problems.

First, while the technical attributes of most PC models were either constant or changed only infrequently, both the list and discount prices changed more frequently, particularly the discount prices. To handle this problem I decided to use the price I found in June as the price for the year. This may distort the estimates of the year-to-year price changes,[18] but seemed the only reasonable way to handle the situation, given the constraints in obtaining data.

Second, on a few occasions there were differences in the list prices for a PC model in a given year. I chose to use prices found in ads over prices found in reviews since the reviews were often done well before a PC was generally available, making the review prices tentative.

Third, to avoid being inundated with discount prices that varied only slightly, I chose to use only the lowest discount price that I found for each model in each year. This may exaggerate the size of the average discount but, since the sources I used are all widespread, anyone looking for a reasonable amount of time would also have been able to find the lowest discount price.

Thus, while some pricing data problems did occur, I have dealt with them so as to prevent any systematic biases from entering into the regression.

3.2 Data Sources

The data for this project were collected from several sources. The primary source was Byte magazine, a journal dedicated to small computer systems. In addition to copious advertising featuring both list and discount prices, it contains comprehensive technical reviews of all new PC systems. I examined every issue of Byte since its inception in July, 1976 to December, 1987, looking in each issue for

---

[18] Cole, p. 44.

the PC reviews. To obtain the pricing data, however, I used only the June issues from each year, since I wished to obtain a mid-year price (as discussed above).

I also used two other computer journals, PC Magazine and PC World, for data collection. I read the 6/23/87, 6/24/86, 6/25/85, and 6/12/84 issues of PC Magazine for pricing data, and read the 1984 and 1985 hardware review issues of PC World to obtain technical specifications on the PCs.

The New York Times proved to be an excellent source for PC discount prices, both within the Science and Technology section and the Business section. The issues I examined were: 6/9/87, 6/10/86, 6/11/85, 6/5/84, 6/7/83, 6/11/82, 6/2/81, 6/2/80. The first PC advertising appeared in 1981, so I did not examine any earlier issues.

I also had access to a sizable amount of IBM pricing and technical data gleaned from IBM marketing announcements and compiled within an earlier paper by the author.[19]

Finally, a personal computer guide from Dataquest[20] was an invaluable source of both pricing and technical data on all the PC models available in 1987. Unfortunately, this was the first year of publication for this guide; otherwise, my data for earlier years might well have been more complete.

3.3 Processing Tools

In doing the analysis for this project I used two software programs. For the collection and sorting of my data I used Lotus 1-2-3 version 2.01.[21] For the

---

[19] Jeremy Cohen, "A Quality-Adjusted Hedonic Price Index for IBM Personal Computers," unpublished paper, MIT Sloan School of Management, Cambridge, MA, 1987.

[20] Dataquest Personal Computer Guide, Dataquest, Inc., November, 1987.

[21] Lotus Development Corporation, copyright 1986, Cambridge, MA.

econometric regression analysis I used Micro-TSP version 5.1.[22] One problem was that Micro-TSP had several restrictions which affected my work.

First, it could only handle 44 variables in a single regression; second, it could handle only 32,000 total data points at any one time. While these restrictions had some effect on my choice of the number of independent variables and on my quick removal of several variables, I do not feel that these restrictions caused any serious problems with my analysis.

3.4 Dependent Variable Description

As described above, the dependent variable in my hedonic price regression is the mid-year price of each PC model for each year. Each of the PC model prices has been deflated to account for the effects of inflation. The average urban Consumer Price Index[23] for June of each year was used to perform the deflation, since PCs are mainly purchased by consumers and businesses for end-use. PCs are usually end-products, as opposed to being inputs into a manufacturing process of some sort.

The CPI data used was:

Year	June CPI (1967=100)
1976	169.2
1977	181.8
1978	195.3
1979	216.9
1980	247.8
1981	271.3
1982	290.6
1983	298.1
1984	310.7
1985	322.3
1986	327.9
1987	340.8

---

[22] Quantitative Micro Software, copyright 1987, Irvine, CA.

[23] Consumer Price Index Detailed Report, U.S. Department of Labor, Bureau of Labor Statistics, Washington, DC.

Another price variable that could have been used was the actual transaction prices for which PCs were sold, as opposed to the advertised list and discount prices. This would give a more unbiased estimate of the hedonic equation,[24] since special quantity discounts, educational discounts and government discounts are quite common. Unfortunately, this information is difficult to obtain not only for current models (since discounting information is often considered proprietary), but even more so for past models. It is also likely that the list prices change less frequently than both the discount schedules and the actual transaction prices, adding a small biasing element into the analysis.

There are a total of 1108 data points in my data set, with each data point consisting of the dependent price variable and the set of independent variables.

3.5 Independent Variable Descriptions

A set of forty independent variables were created for the right-hand side of the regression equation. These variables are broken down into two different groups. The first group of fourteen is continuous, each variable representing some measurable physical attribute (quality factor) of a PC model at a given point in time.

The second set of twenty-six are dummy (0 or 1) variables that are grouped into the following eight subdivisions:

1) Processor type
2) Monitor type
3) Portability
4) Additional technical features
5) Price type
6) Manufacturer
7) Year of sale
8) Pre/Post IBM date

[24] Cole, p. 44.

The dummy variables within each subdivision span the universe of possibilities for one feature or aspect of the PCs.

The forty initial independent variables are described in the following sections.

3.5.1 Continuous Variables

1) **RAM** - The amount of Random Access Memory (RAM) standard on each PC model, measured in K (1024) bytes. This information was obtained for all 1108 data points.

2) **Maximum RAM** - The maximum amount of RAM that can fit on the system board (mother-board) of each PC model, measured in K bytes (logical groupings of eight on/off bits). This figure does not include possible RAM increases due to expansion cards, since RAM access is typically faster on the system board than on expansion cards, placing a premium on the amount of fast system board RAM that the system can have. Also, expansion slots are typically multi-functional and not restricted to just RAM cards, so the amount of expansion card based RAM can be quite variable for most PC models. This information was obtained for all 1108 data points.

3) **ROM** - The amount of Read Only Memory (ROM) standard in each PC model, measured in K bytes. Since ROM usually contains diagnostics and low-level operating system routines, this variable may well serve as a surrogate for the sophistication of the software packages in the PC. This information was difficult to obtain for many of the older PCs, resulting in only 325 entries among the 1108 data points.

4) **Mhz** - The clock speed of each PC mode, measured in Megahertz (million of cycles per second). This is one of the main indicators (along with processor type) of the throughput of a PC. This information was obtained for all 1108 data points.

5) **Hard Disk** - The amount of storage on the hard disk (if one exists) in each PC, measured in M (1024*1024) bytes. A hard disk is often the dividing line be-

tween a home system and an office system, though this distinction has blurred in recent years. 332 of the 1108 models in the data set had hard disks.

6) **Hard Disk Access Speed** - The average time it takes to retrieve a byte of information from the hard disk, measured in milliseconds. This information would be more useful if it were supplemented by the seek and rotation times of the hard disks, but it was extremely difficult to obtain this data. It was also somewhat difficult to get the access speed data, with it being found for only 107 of the 332 PCs with hard disks.

7) **Floppy Disk** - The amount of storage that the floppy disk drives, if any exist, are capable of reading or writing to a floppy disk, measured in K bytes. This includes the flexible 8″, 5.25″ and 3.5″ media and excludes fixed media. Of the 1108 PC model data points, 898 had one or more floppy disk drives.

8) **Number of Floppy Disk Drives** - The number of floppy disk drives standard on each PC model. A variable equivalent to this one for hard disk drives is not necessary, since all of the PC models examined had either zero or one hard disks, while the PCs had either zero, one, or two floppy disk drives. This data was obtained for all 1108 data points.

9) **Slots (8 bit)** - The number of eight bit expansion slots available within each PC model for expansion boards. This data was obtained for all 1108 data points.

10) **Slots (16 bit)** - The number of sixteen bit slots available within each PC model for expansion boards. This data was obtained for all 1108 data points.

11) **Slots (32 bit)** - The number of thirty-two bit slots available within each PC model for expansion boards. This data was obtained for all 1108 data points.

Note: The above three "slot" variables posed a few problems. First, both PC ads and reviews often failed to state whether any of the slots mentioned were already filled within the standard setup. To be consistent I have used the total number of slots in the system, irrespective of whether the slots may have been initially filled. This may not be optimal, but it was the best solution given the available data. Since a machine would normally not have more than one or two slots initially filled, this should not cause any major problems.

Another problem is that some ads and reviews specified only the number of slots, not the size of the slots. I resolved this problem through the use of multiple data sources (particularly the Dataquest guide). In the few cases where the situation was still unresolved, I assumed that the size of the slots was the same as the size of the PC's processor chip. This is a reasonably safe assumption that should not have any negative effects on the regression analysis.

12) **Size** - The size of each PC model, measured in cubic inches. This includes the system unit but normally excludes the weight of the monitor or keyboard. These components are included, however, if they are integral to the system unit (i.e., IBM Convertible). This data was obtained for 757 of the 1108 data points.

13) **Weight** - The weight of each PC model, measured in pounds. This includes the system unit but normally does not include the weight of the monitor or keyboard. These components are included, however, if they are integral to the system unit (i.e., Apple Macintosh). This data was obtained for 655 of the 1108 data points.

14) **Age** - The number of years since a given PC model was first introduced. A model has an age of zero its initial year. This variable ranges from zero to seven (for the Radio Shack Color Computer) and provides useful information on the effects of longevity on pricing.[25] This data was obtained for all 1108 data points.

[25] Note that while the specifications of many PC models changed over time, as long as the model name remained constant the model was considered to be the same as the model from the previous year.

Age	Number of Data Points
0	649
1	257
2	118
3	46
4	25
5	9
6	3
7	1
	1108

Note: The hard disk size, number of floppy disk drives, number of slots, and age variables had the value one added to them so that it would be possible to take their natural log during the regression analysis.

3.5.2 Dummy Variables

The twenty-six dummy variables are divided into eight subdivisions, as described above. The variables and subdivisions are:

I) Processor Type - All of the PCs in my study have either eight bit, sixteen bit or thirty-two bit processors, this being an indication of how much data the system can process at a given time. The higher this number, ceteris paribus, the greater the throughput of the system.

A few processor chips, such as the 68000, can manipulate differing amounts of data depending on the operation. In this case I have grouped such multiple-size chips in the lower applicable group, both because this is how these chips are normally viewed and because the throughput of the chip is restricted by its lowest grouping.

15) **DProc16** - 1 if the system has a sixteen bit processor chip, 0 otherwise.
16) **DProc32** - 1 if the system has a thirty-two bit processor chip, 0 otherwise.

Thus, the base case for this subdivision is having an eight bit processor chip. In my 1108 data point sample, 540 of the PCs had eight bit processor chips, 506 had sixteen bit processor chips, and 62 had thirty-two bit processor chips.

II) Monitor Type - While many PCs are sold without a monitor, they are also sometimes sold either with a black and white (B&W) or a color monitor.

17) **DBW** - 1 if the system comes with a B&W monitor, 0 otherwise.
18) **DColor** - 1 if the system comes with a color monitor, 0 otherwise.

Thus, the base case for this subdivision is not having a monitor. In my 1108 data point set, 605 of the PCs had no monitor, 478 had a B&W monitor and 25 had a color monitor.

III) Portability - Some PCs, often called ″portables″ or ″convertibles,″ are made small and light enough to be portable. These PCs often also have special features such as battery power capability and an integral monitor.

19) **DPortable** - 1 if the system is meant to be portable, 0 otherwise.

Thus, the base case for this subdivision is not being portable. In my 1108 point data set, 937 of the systems were not portable, while 171 were explicitly portable.

IV) Additional Technical Features - Some PCs have extra hardware that is costly enough to have a significant effect on their overall price, yet rare enough not to be considered a standard item. Some examples include modems, printers, or an extra monitor.

20) **DExtra** - 1 if the system has a significant piece of additional hardware, 0 otherwise.

Thus, the base case for this subdivision is not having any additional equipment. In my sample set, only 24 of the 1108 systems had additional equipment of this sort.

V) Price Type - In my sample I have both list prices and discount prices. This allows an analysis of discount pricing and also provides more overall variability in the pricing data. In particular, it is not unusual to see a PC's list price remain steady over 2-3 years while its discount price may drop steadily over this time.

21) **DDiscount** - 1 if the system price is discounted, 0 otherwise.

The base case for this subdivision is having a list price. In my sample, 841 of the systems had list price information and 267 had discount prices.

VI) Manufacturer - In my data set I have PCs from 114 different companies. In this subdivision, I attempt to discover any differences in pricing among seven major PC manufacturers. While any pricing discrepancies may possibly be accounted for by intangibles such as quality and reliability or more tangible items such as warranties and included software, these discrepancies may also be an indicator of a company's overall pricing policy.

22) **DApple** - 1 if the PC is made by Apple, 0 otherwise.
23) **DCommo** - 1 if the PC is made by Commodore, 0 otherwise.
24) **DCompa** - 1 if the PC is made by Compaq, 0 otherwise.
25) **DIBM** - 1 if the PC is made by IBM, 0 otherwise.
26) **DNEC** - 1 if the PC is made by NEC, 0 otherwise.
27) **DPCLim** - 1 if the PC is made by PC Limited, 0 otherwise.
28) **DRadio** - 1 if the PC is made by Radio Shack, 0 otherwise.

The base case for this subdivision is to be manufactured by one of the 107 other PC companies. Of the 1108 PC data points, 62 were made by Apple, 40 by Commodore, 59 by Compaq, 94 by IBM, 36 by NEC, 21 by PC Limited, 85 by Radio Shack, and 711 by other companies. Thus, 35.4% (396 of 1108) of the models were built by these seven manufacturers.

VII) Date - The heart of a hedonic pricing study are the yearly dummy variables. The data in my sample runs from 1976 to 1987, resulting in eleven of these dummy variables. The parameter coefficients obtained for these variables will be directly used to construct the hedonic price index.

29) **D77** - 1 if this model/price data point is from 1977, 0 otherwise.
30) **D78** - 1 if this model/price data point is from 1978, 0 otherwise.
31) **D79** - 1 if this model/price data point is from 1979, 0 otherwise.
32) **D80** - 1 if this model/price data point is from 1980, 0 otherwise.
33) **D81** - 1 if this model/price data point is from 1981, 0 otherwise.
34) **D82** - 1 if this model/price data point is from 1982, 0 otherwise.
35) **D83** - 1 if this model/price data point is from 1983, 0 otherwise.
36) **D84** - 1 if this model/price data point is from 1984, 0 otherwise.
37) **D85** - 1 if this model/price data point is from 1985, 0 otherwise.
38) **D86** - 1 if this model/price data point is from 1986, 0 otherwise.
39) **D87** - 1 if this model/price data point is from 1987, 0 otherwise.

The base case for this subdivision is a model/price data point from 1976. The number of data points from each year is as follows:

Year	# of Data Points		
1976	11		
1977	17		
1978	15		
1979	28		
1980	45		
1981	40	156	(1976-1981)
1982	54		
1983	85		
1984	130		
1985	109		
1986	178		
1987	396	952	(1982-1987)
	1108		

VIII) Pre/Post IBM Date - IBM revolutionized the PC industry when it released its PC line at the end of 1981. Since this may well have affected overall industry pricing, I examined it using the following dummy variable:

40) **DP82** - 1 if the system data point is from after 1981, 0 otherwise.

The base case for this subdivision is for the data point to be from before 1982. Of the 1108 data points, 156 are from before 1982, while 952 are from 1982 or later.

3.6 Overall Base Case

The overall base case (in which all of the dummy variables equal 0) is: a PC model with an eight bit processor, no monitor, non-portable, with no extra hardware features, list price, not from one of the seven specified manufacturers, from 1976, and from before the late 1981 IBM announcement date for its PCs.

3.7 Omitted Variables

There were several independent variables that I considered using in the initial regression equation but eventually chose not to use, for various reasons. These variables are:

1) Hard Disk Seek Speed - This information proved to be very difficult to obtain for the early PC models, since it was rarely reported either in advertisements or in product reviews. However, it should prove to be highly correlated with the hard disk access speed variable, which is included in the regression. Thus, I conjecture that this variable would not have added much new information to the analysis.

2) Floppy Disk Access/Seek Speed - This information was difficult to obtain for recent PC models and often impossible for the earlier models. Thus, I was not able to gather enough entries for this variable to make it worthwhile for inclusion in the regression.

3) Asynchronous (Asynch.) Card - While the information on whether each PC came with an Asynch. card was often available, I decided that this variable would likely not add much to the regression analysis, since the cost/value of an asynch. card is quite low (on the order of $20-40).

4) Clock/Calendar Card - Similar to the asynch. card, the low cost/value of a clock/calendar card (about $20) implied that it was not worth adding it to the regression, even though this information was often available.

5) Floppy Disk Size - While the information was available on whether the floppy disks used by the system were 8″, 5.25″ or 3.5″, I felt that this variable would be highly correlated with both the Floppy Disk Size variable and the Year variables (since 3.5″ disk drives are quite new). Thus, I conjecture this variable would also not have added much information to the regression.

6) First Offered - I had intended to have a dummy variable specifying when each PC model was first offered. This variable would have helped to explore the possibility of differential pricing when a PC model is introduced. Instead of using this variable, however, I chose to use the Age variable that is included in the initial variable list. The Age variable not only provides the information that the First Offered variable would have provided, but it also gives additional information concerning the effect of PC model longevity on its price. Since the First Offered variable is subsumed by the Age variable, it would not have added any additional information to the regression.

7) Top-of-Line - This would have been a dummy variable indicating that a given PC model was the best one at that time from a given manufacturer. This would have allowed an examination of any premium pricing that may have been going on within the industry. Unfortunately, the PC lines of most manufacturers changed often and sometimes had several models at or near the top, making it impossible to determine which models were indeed top-of-line. Thus, this variable had to be omitted.

8) Bottom-of-Line - Similar to the above variable, this variable would have indicated that a given PC model was at the bottom of the line for a given manufacturer in a given year. This would have helped in assessing whether there were any low end pricing effects. Unfortunately, I was unable to determine which models were bottom-of-line, so this variable had to be eliminated.

9) Processor Type - Instead of starting with the broad groupings of eight bit vs. sixteen bit vs. thirty-two bit that I eventually accepted, I considered using dummy variables to indicate the processor chip being used (for example 8086, 80286, 68000, etc.). I chose not to do this for two reasons.

First, it would have necessitated the use of at least 10-12 additional dummy variables, forcing the omission of other variables. Second, while all sixteen bit processors are not created equal, they are usually quite similar.

Thus, between the eight, sixteen, and thirty-two bit dummy variables and the Mhz variable, most of the information that the processor type dummy variables would have provided is already available. Therefore, the processor type dummy variables were omitted.

10) Final Assembly Region - This would have been a series of 8-10 dummy variables indicating where each model underwent final assembly (for example, Japan, US, Taiwan). While providing interesting information, the cost in the number of dummy variables necessary was high and the data was only readily available for very recent PCs. Thus, these variables were eliminated.

11) Operating System - This would have been a series of 4-5 dummy variables indicating which operating system(s) each PC model could run (for example, CP/M, MS-DOS). This would have provided an interesting tie-in between software value and pricing, but was omitted for two reasons.

First, the great majority of recent systems (95+%) run some version of MS-DOS, so these variables would not have provided much differentiation. Second, while the information is readily available for recent systems, it would have been much more difficult to obtain for the older models.

Thus, these variables were left out.

12) RAM Access Speed - While the RAM access speed is one of several indicators used to describe the throughput of a PC, it is normally highly correlated with

the clock speed (Mhz) of the system. Thus, this variable would not have added much new information to the regression and it would also have been quite difficult obtaining this data for PCs before 1985. For these reasons this variable was not used.

13) Number of Wait States - While the clock speed is one indicator of a PC's throughput, it can often be modified significantly by the number of wait states used when accessing RAM. Unfortunately, it was impossible to find this information for any but the most recent PC models, so I decided to omit the variable. While this does cause some loss of precision in the analysis, the Mhz and Processor Size variables should still explain most of the effects of PC throughput on pricing.

4.0 Procedure and Results

In this section I discuss the techniques used in the derivation of my hedonic price index for personal computers, as well as the results I obtained.

4.1 Choice of Equation Form

The three types of equation form most often used for hedonic price indexes are log-log, semi-log, and pure linear. To choose among these three possibilities, I first plotted some of the independent variables versus the dependent variable in all three formats. The linear plots appeared to be slightly better than the other two formulations, but the difference was marginal.

Next, I examined the correlation coefficients of the variables in all three forms. The linear form independent variables were indeed more highly correlated with the dependent variable than the other forms, but there was also a much greater multicollinearity effect. While multicollinearity can have positive effect on our hypothesis testing, as we shall see in section 4.3, it is normally best to avoid it, if possible.

In a review of the literature, I discovered that not only was the log-log form the most commonly used form for hedonic price indexes,[26] but specific studies on computer peripherals[27] and mainframe computer systems[28] found this form to give the best results using the Box-Cox formulation.

The log-log form is also easier to work with than the other two equation forms since the yearly dummy variable coefficients can be directly transformed into the quality-adjusted price indexes by taking their antilogarithm. This conversion is not

---

[26] Triplett, Table 2.

[27] **Ibid.**, p. 34.

[28] Cole, p. 44.

nearly as mechanistically straightforward in the other two forms.

Finally, from a theoretical point of view, the log-log form makes more sense than the linear form since we would expect that the law of diminishing marginal utility would apply to this situation. For example, I expect that doubling the values of the quality variables would result in a less than doubling of the price, which fits better with the log equation forms than the linear form.

The log-log form assumes a constant elasticity and a varying slope, while the linear form assumes a varying elasticity and a constant slope. In the rapidly changing PC market, the former conditions seem more applicable.

Thus, based on the above arguments, I have decided to use a log-log form for my hedonic price equation.

4.2 Data Restrictions

After transferring my data from Lotus to Micro-TSP, I began to run some preliminary correlations and regression analyses. I quickly discovered, however, that in my first regression using 500 data entries only 16 of the entries were used. I determined that the cause of this problem was that only 16 of my 500 data points contained entries for all 40 of the independent variables.

I attempted to extrapolate values for some of the missing data but did not feel comfortable doing this, especially for older systems with which I was not personally experienced. Thus, I decided temporarily to remove the independent variables that were causing most of the problem. These variables were ROM, Hard Disk Access Speed, Size, and Weight. Removing these four variables resulted in 97% (1076 of 1108) of my data points being utilized in my subsequent analyses.

The effects of these four temporarily removed variables will be re-examined in section 4.6, after the initial analysis of the data has been completed. Removing these four variables resulted in the number of independent variables in the regression equation dropping from forty to thirty-six.

4.3 Multicollinearity

Multicollinearity occurs when two or more of the independent variables in the regression equation are highly correlated with one another. Two variables are highly correlated if a large amount of the variation in one variable is reflected in a corresponding variation in the other variable.

Normally, multicollinearity increases both the true and estimated standard errors of the affected variables. As the absolute value of the correlation coefficient between two independent variables increases, it becomes more difficult to reject the null hypothesis that one of their regression coefficients is zero. The resulting decrease in the variable t-statistics reduces our ability to analyze the effect of each independent variable. The estimated parameter might also change, so it is indeterminate as to what the final effect will be on the t-statistic, although in practice it is more likely that it will be decreased.

Multicollinearity can also have a positive effect, however. Since it reduces the variance of the sum or difference between coefficients, it can enhance the precision with which certain hypotheses can be tested. For example, if variables 1 and 2 have high positive correlation, it is possible to make a superior test of $B_1 + B_2$, where B_1 and B_2 are the estimated regression coefficients for variables 1 and 2, respectively.; if they have high negative correlation, it is possible to make a superior test of $B_1 - B_2$.

Thus, multicollinearity can have both positive and negative effects. However, for my study it proved to be a problem for two reasons. First, the type of hypothesis testing I am most interested in is the $B_i = 0$ type, which is negatively affected by multicollinearity. Second, when I initially tried to run a global regression, Micro-TSP was unable to perform the calculation because of excessive multicollinearity between the independent variables.

Having decided that some of my independent variables were too highly correlated to be useful, I ran correlations between each pair of variables to find the "culprits." I used the rule of thumb that, if the correlation between two independent variables is larger than the correlation of either with the dependent variable then a problem might exist.[29] I also assumed that any correlation above .8 was suspect.

Using these guidelines, I narrowed my search for excessively correlated independent variables to Floppy Disk Size, Number of Floppies, Maximum RAM, RAM, and Mhz.

Examining this group, I saw that Floppy and NumFloppy were excessively correlated (85%), as were MaxRAM and RAM (88%). I decided to remove the Floppy and MaxRAM variables because I felt they were the least important of each pair. I also combined the Slot8, Slot16, and Slot32 variables into one Slots variable, since the Slot16 variable was highly correlated (70%) with the DProc16 variable and the Slot32 variable was highly correlated (90%) with the DProc32 variable.

Another variable I withdrew was DP82, the pre/post IBM PC date dummy variable. While it was not correlated at higher than 80% with any other one variable, it was correlated at the 60-70% level with 4-5 other variables, preventing the regression from being resolved.

I left the Mhz variable in the equation even though it was correlated with RAM, since I felt that it was one of the main measures of processor throughput and therefore an important independent variable.

---

[29] Pindyck, p. 89.

The final result of my investigation into the effects of multicollinearity was the reduction in the number of independent variables from thirty-six to thirty-one. After doing this, I was able, for the first time, to run a regression, whose results are given in Appendix A.

4.4 Reducing the Regression Equation

While the regression equation in Appendix A explains about three-quarters of the variability in the dependent variable, I felt that a few of the independent variables in the equation were not contributing much, if anything, to the explanatory value of the regression. The next step in analyzing the data was to see what effect removing these variables had on the equation.

The initial equation F-statistic was 93.19, which is much higher than the $F_{31,930}$ 5% critical value of 1.47. This indicates quite strongly that the independent variables are indeed not jointly equal to zero.

Upon examining the t-statistics of the independent variables, however, it was apparent that the DNEC and DRadio variables were not contributing much to the regression. The t-statistic for DNEC was .033, while for DRadio it was -.856. Each of these t-statistics have absolute values well less than 2, so at a 5% confidence level each is insignificant.

Thus, I removed these two variables from the regression and then performed an F-test of the hypothesis that they were jointly equal to zero. The $F_{2,930}$ value I obtained was .37, which is much less than the 5% critical value of 3.01. This F-test result indicates that removing DNEC and DRadio is certainly appropriate.

Next, I examined the regression equation again for any other variables with an absolute value t-statistic less than 2. Both DIBM (-1.315) and DApple (1.894) fit this description, so I decided to test them further.

After removing DIBM from the regression equation I obtained an $F_{3,930}$ value on removing DNEC, DRadio, and DIBM of .82, still less than the 5% critical value of 2.61. Since DIBM "failed" both the F- and t-tests, I decided to remove it from the equation.

I now looked at the DApple variable and saw that its t-statistic value had risen from 1.68 originally to 2.05, which indicated that the variable was now significantly different from zero. I tried removing it and obtained a four variable F-statistic value of 1.68. This is still less than the $F_{4,930}$ 5% critical value of 2.38, but it is substantially closer. Because of this and because of its increased t-statistic value, I decided to leave DApple in the equation.

Finally, I performed one last joint F-test on DApple and DColor, since they had the least significant t-statistics. The $F_{2,930}$ value of 4.31 that I obtained was greater than the 5% critical value of 3.01, so both the F- and t-tests indicate that these two variables are significant and therefore belong in the equation.

Note that throughout this process I have used a mixture of both "science" and "art" to make my decisions on variables. There are no hard and fast rules to this process. I have instead used my judgment based on the information I received from the various regression results and variable tests. The main guideline I have followed is that, ceteris paribus, the fewer explanatory variables the better.

Thus, in this section I decided to remove DNEC, DRadio, and DIBM, reducing the number of independent variables from thirty-one to twenty-eight. This reduction allowed me to use about 100 more data points in the analysis, due to Micro-TSP restrictions, resulting in the regression equation shown in Appendix B.

4.5 Potential Data Problems

Whenever regression analysis is used to examine data, several potential data problems can create pitfalls for the unwary. This section explores two of the most common of these data problems and examines how much difficulty they posed for me.

4.5.1 Autocorrelation

First-order autocorrelation occurs when the error terms from adjacent observations are correlated. This is often the case in time-series studies, either because of similar measurement errors or because of the high degree of correlation over time present in the cumulative effects of omitted variables in the regression model.[30]

For my model I would not expect first-order autocorrelation to be a problem, since little measurement error should exist and because the data is more of a cross-sectional nature rather than time-series.

There is also no reason to suspect that higher order autocorrelation would be a problem, both for the above reasons and due to a lack of items such as seasonal lags in the data. Thus, autocorrelation is not a concern for my regression analysis.

4.5.2 Heteroscedasticity

Heteroscedasticity occurs when the error terms in the regression have unequal variances. This does not usually occur in time-series studies,[31] but my data set has enough cross-sectional aspects for this problem to be potentially significant.

To test for this I first ran the Goldfeld-Quant test on several of the variables in

---

[30] Ibid., p. 153.

[31] Ibid., p. 141.

the data set and received positive results for a few. The most significant one was the dependent variable LRPrice. It had a test statistic of 2.35, which was greater than the $F_{397,397}$ (N = 1055, d = 211, k = 25) 5% critical value of 1.05.

To check this result I ran a regression of LRPrice on the residuals (from Appendix B) squared and obtained a t-statistic of -7.2, indicating that the larger the log of the real price of the PC, the smaller the error term. LRPrice explained only about 4% of the total variation in the error terms. Therefore, the problem, while existing, did not seem to be making a major effect on the results.

To correct the apparent problem I weighted all of the data by one over the square root of the fitted values for LRPrice (LRPrice - residuals) and then estimated the transformed data by ordinary least squares. This is called generalized least squares, and it resulted in the regression equation in Appendix C. As can be seen, very few of the parameter coefficients, standard errors, or t-statistics have changed more than 10% due to the weighting. This was also the case when the weighting was done using several other variables, including RAM and Mhz. It appears that while some heteroscedasticity is present, its effects are minimal.

Thus, for the sake of simplicity, I decided to leave an in-depth analysis of this problem for a future study. This will undoubtedly leave some minor bias in the standard errors of the parameters, but the amount of bias seems quite small and should not at all affect the conclusions I draw from the hedonic regression results.

4.6 Review of Temporarily Deleted Variables

In section 4.2 the variables ROM, hard disk access speed, size, and weight were removed from the regression equation due to the relatively small number of entries for which this data was available. This section examines the impact of each of these variables on the reduced equation developed in section 4.4.

4.6.1 ROM

When the ROM variable is added to the equation, only entries from 1983 and later can be used since no ROM data is available before then. The effect of adding this variable is that the coefficient and t-stat of the RAM variable drop substantially (coef. = .1677, t = 4.292) and the ROM variable is quite significant (coef. = .103, t = 3.9).

It appears that RAM and ROM are strongly correlated and that the RAM variable is acting in part as a placeholder for ROM. When ROM is added, however, both ROM and RAM are positive and significant. RAM seems to have a greater impact, but ROM is also quite significant.

4.6.2 Hard Disk Access Speed

When this variable is added to the regression equation, Micro-TSP is unable to resolve the regression due to multicollinearity problems. The access speed is highly correlated with the size of the hard disk (corr. = -.746) and with the Mhz of the system (corr. = -.678).

The larger the hard disk and the faster the system, the faster the access speed of the hard disk, which is what one would expect.

4.6.3 Size

Adding the Size variable to the regression equation results in a noticeable change. The coefficient and t-statistic for size are very high (c=.279, t=9.07), indicating that the larger the system, the more expensive it is. This does not seem to make sense, except that a large system would be more likely to have components such as disk drives, hard disks, or monitors that would add to its value.

Interestingly, the DPortable variable now becomes more significant in the regression (c=.502, t=9.59), indicating that a smaller size is indeed of some value.

4.6.4 Weight

Adding the Weight variable also causes a change in the results. The coefficient and t-stat are high (c=.410, t=10.15), indicating that the heavier a system is, the more expensive it is. This appears counter-intuitive, but a greater weight might represent monitors (DBW and DColor have t-stats of 1 and 1.5, respectively), hard disks, floppy disks, as well as sturdier construction.

The Size and Weight variables seem to be representing indirect effects, i.e., they do not have a direct effect on a PC's price, except as noted above. PCs with additional features are likely to be bigger and heavier, but it is the features themselves that add to the price, not their size or weight.

Thus, adding either the Size or Weight variables to the regression equation does not seem appropriate. Neither does adding (Size/Weight), which gave results consistent with both of the individual variables.

Also, since the addition of these independent variables had little effect on the time dummy variables, their presence has little effect on the resulting hedonic price index.

4.7 Overall Regression Equation

The regression equation in Appendix B represents my preferred explanation of the variation of the log of the real price of the PCs. In this section I examine the coefficients and t-statistics for each of the independent variables in terms of relative importance and significance. Most of the results are relatively straightforward;

however, a few of the coefficients are particularly interesting. The relative importance of the dummy variables is also accessed.

The following table shows the independent variables (except for the year dummies, to be reviewed in section 4.10), in order of their statistical significance.

Independent Variable	t-statistic
RAM	16.38
Hard Disk	15.52
Slots	9.20
Num. Floppies	8.86
Discount	-8.83
Portable	5.99
PC Limited	5.61
Commodore	-5.39
Mhz	5.16
Processor - 16 bit	5.13
Processor - 32 bit	4.73
B&W Monitor	4.56
Compaq	4.45
Extra Equip.	3.21
Age	2.98
Apple	2.48
Color Monitor	2.22

The first five independent variables are noticeably more significant than the rest. Using only these five variables, an R^2 of .433 was obtained, indicating that these variables provide a large part of the explanatory power of the model. With all the independent variables included, the R^2 was .746.

Each of the variables is significant at the 95% confidence level and each except Apple (98.6%) and Color Monitor (98%) are significant at the 99+% confidence level.

The dummy variables can be assessed in a number of ways. Two ways of special interest are the value of the estimated coefficient and the implied effect on a $4,000 PC. Both are displayed in the following table.

Dummy Variable	Coefficient	Effect on $4,000 PC
Processor - 32 bit	.385	$1,878
Extra Equip.	.294	$1,367
Compaq	.282	$1,303
Portable	.263	$1,203
Color Monitor	.203	$ 900
Processor - 16 bit	.191	$ 842
Apple	.159	$ 689
B&W Monitor	.138	$ 592
Discount	-.299	-$1,034
Commodore	-.383	-$1,273
PC Limited	-.547	-$1,685

Thus, having a 32-bit processor adds the most to a PC's price, closely followed by having an extra component and by being a Compaq system. At the other end, PC Limited systems, Commodore systems, and discounted systems are below the average PC price.

There are several relationships among the dummy variables that add to the credibility of the results. For example, a 32-bit processor adds more to the price of a PC than does a 16-bit processor. Also, having a color monitor adds more than does having a B&W monitor.

4.8 Independent Variable Parameter Results

In this section the results obtained for each of the seventeen independent variables are reviewed. As mentioned in the previous section, each of these variables is significant at the 99% confidence level, except for DColor and DApple.

The yearly dummy variables are not reviewed here but are instead examined separately in section 4.10.

RAM - The highly significant positive coefficient indicates both that the amount of RAM is a major determinant in the price of a PC and that the greater the amount of RAM in a PC, the higher its price. This is indeed the result that I expected.

Mhz - The significant positive coefficient indicates that the higher the clock speed of the system, the higher its price. This suggests that the greater the power (throughput) of a PC, the more expensive it is, a reasonable result.

Hard Disk - The highly significant positive coefficient indicates that having a hard disk adds substantially to the price of a PC. This effect increases as the size of the hard disk increases, which is also to be expected.

Number of Floppies - The number of floppy disk drives that a PC has is also a highly significant determinant of its overall price, as one would expect. The more floppy disk drives a PC has, the higher its price.

Slots - The total number of slots that a PC has for expansion boards has a highly significant positive effect on price. While the direction of this result is expected, it is somewhat surprising that the coefficient is so large. It appears that PC users and therefore PC manufacturers place a sizable premium on the ability to expand their PCs, even though the slots themselves provide no inherent functionality.

Age - This variable also produces surprising results. The positive coefficient indicates that the older a PC system, the more expensive it will be, relative to newer models. Although this does not make intuitive sense, there are a few possible explanations for this effect.

First, while empirically one does see the prices of old PC models drop as new models appear, the price drop may not be large enough to put the older model at an equivalent price to the new model. Manufacturers no doubt realize that even though old models will rapidly become obsolete, the existence of switching costs and product loyalty may result in significant sales of older models, even when newer and relatively less expensive models are available. One example of this might be the Apple IIE, which has continued to sell well even though obsolete, due to its low absolute (not relative) price and the huge install base of software compatible with it.

Another explanation of why older models may not be dropped in price "enough" when new models appear is that there is often a limited supply of the new models and some purchasers prefer to buy the old model now rather than wait for the new model. Also, it can sometimes take up to a year before the public becomes fully aware of a new model. Thus, time lag effects may be apparent as well.

Finally, the rapid changes in the PC marketplace, including all the technological "leapfrogging," may well mean that the market is never at a long-term equilibrium.[32]

When a significantly enhanced model appears many of the older models disappear within a year. The older models that do remain are likely to be the ones that have some competitive advantage or niche allowing them to compete with the newer, more sophisticated and relatively less expensive models. Thus, while the prices of the old models drop, they may not need to drop to equilibrium with the newer mod-

[32] Triplett, p. 16

els. If true, this could help explain why the Age variable has a positive coefficient.

16-Bit Processor - A system with a 16-bit processor is significantly more expensive than a system with an 8-bit processor, as one would expect. The added throughput results in a higher price for the PC.

32-Bit Processor - A system with a 32-bit processor is considerably more expensive than a system with either an 8-bit or 16-bit processor, which is reasonable. The additional throughput results in a much greater price for the PC.

Black & White Monitor - A PC that comes with a B&W monitor is significantly more expensive than a PC that does not come with a monitor. This is reasonable since a B&W monitor would normally add between $100-200 to the price of a PC system.

Color Monitor - A PC that comes with a color monitor is much more expensive than a PC that either has no monitor at all or else has only a B&W monitor. This is not surprising, since a color monitor would usually add about $200-500 to the price of a PC.

Portable - The highly significant positive coefficient indicates that portable PCs are priced noticeably higher than the equivalent non-portable systems. This is what one would expect due to the extra costs of designing and producing portable computers and the additional value they provide. They are also typically sold to a less price-sensitive market, which may explain some of their price premium.

Extra - A PC that has additional features such as modems or printers will be significantly more expensive than a PC without these extra features. This is, of course, an expected result.

Discount - A PC that is being sold by a wholesaler or retailer at a discount will be significantly less expensive than the same system bought from the manufacturer at list price. The highly significant negative coefficient indicates that the average discount from the list price is quite sizeable. Most of the PCs are available at a discount from one place or another, except for PCs from a few manufacturers such as PC Limited, who sell their systems themselves only at list price.

Apple - PCs from Apple are significantly more expensive than equivalent systems from other manufacturers. It is difficult to tell, however, whether this pricing differential is due either to quality attributes omitted from the regression equation, or by an overall premium pricing policy set by the manufacturer. Since Apple has the reputation as a premium quality PC manufacturer, both explanations may be at least partially correct.

Commodore - PCs from Commodore are considerably less expensive than systems from other manufacturers. This seems to indicate that Commodore is pricing aggressively, since its systems are roughly comparable to many of the PC offerings in the marketplace.

Compaq - PCs from Compaq are significantly more expensive than PCs from other manufacturers, including Apple. Compaq is widely viewed as the hardware technology leader in the PC market and is often the first to introduce new technology such as the 80386 microcomputer chip. Thus, their higher prices probably represent a premium for their technology and market leadership.

PC Limited - PCs from PC Limited are significantly less expensive than other PCs. PC Limited is an IBM PC clone company that advertises and relies on very low prices to provide it with a competitive advantage in the PC market.

4.9 PC System Price Breakdown

In this section the relative effect of each of the independent variables is compared through the use of a hypothetical PC. This "vanilla" PC is not made by a major manufacturer, sells at list price, is not portable, has no extra features, does not come with a monitor, has an 8-bit processor, has been on the market for one year, has 6 expansion slots, 1 floppy disk drive, a 20M hard disk, runs at 4.77 Mhz, and comes with 640K RAM.

The estimated price of this PC in 1976 is computed as follows:

Attribute	Coefficient*LN(attribute)
constant	6.28
RAM	2.40
Mhz	.31
Hard Disk	.54
Floppy Disk	.31
Slots	.36
	10.20 = **$26,903**

Thus, the estimated price for this system in 1976 is $26,903. To obtain a more more empirical view of the relative importance of the independent variables, I show in the following chart how the overall price of the above system would change, given individual changes in the quality attributes listed below:

Attribute	Amount	Coeff*LN(attr)	New Price
RAM	256K	2.06	$19,149
RAM	1024K	2.58	$32,209
Mhz	8	.42	$30,031
Mhz	12	.50	$32,533
Hard Disk	0M	0.00	$15,678
Hard Disk	40M	.66	$30,333
Floppy Disk	2	.49	$32,209
Slots	0	0.00	$18,770
Age	3yrs	.13	$30,638
16-bit proc.	1	.19	$32,533
32-bit proc.	1	.39	$39,735
B&W monitor	1	.14	$30,946
Color monitor	1	.20	$32,860
Portable	-	.27	$35,242
Extra	-	.29	$35,954
Discount	-	-.30	$19,930
Apple	-	.16	$31,571
Commodore	-	-.38	$18,398
Compaq	-	.28	$35,596

PC Limited	-	-.55	$15,522
1977	-	-.59	$14,913
1978	-	-.84	$11,614
1979	-	-.98	$10,097
1980	-	-1.01	$ 9,799
1981	-	-1.18	$ 8,267
1982	-	-1.49	$ 6,063
1983	-	-1.99	$ 3,677
1984	-	-2.05	$ 3,463
1985	-	-2.48	$ 2,253
1986	-	-2.91	$ 1,466
1987	-	-3.21	$ 1,086

While many of the attribute independent variables such as RAM, Hard Disk, Slots, and Processor Size have major effects on the PC system price, the "descriptional" variables such as Discount and Manufacturer are also quite important. However, the effects of both these groups of variables are dwarfed by the effects of changes in the yearly dummy variables.

As the chart shows, the original 1976 price of $26,903 for the base system drops to $1,086 by 1987. This is a precipitous drop in price, indicating that dynamic changes have been occurring in the PC market over the last 11 years.

The effects and meaning of these yearly dummy variables are further explored in the next section.

4.10 <u>Hedonic Price Index</u>

In this section the coefficients of the yearly dummy variables are used to create an hedonic price index for the 1976 through 1987 time period. The coefficients are then examined individually and jointly for statistical significance.

As can be seen in Appendix B, all of the yearly dummy variables have t-statistics with an absolute value greater than 3.4, ranging up to 18.8. This indicates that each of the yearly dummy variables is individually significant at the 99% confidence level. In addition, a joint $F_{11,1024}$ test resulted in a test value of 78.22, far greater than the 5% critical value of 1.8. This means that the yearly dummy variables are not jointly equal to zero; i.e., real personal computer prices have been changing significantly over this time period.

The yearly dummy parameters, the yearly price index (computed by taking the anti-log of each coefficient) and the resulting yearly real price decreases are reproduced in column one of Appendix D. The price index is for a "vanilla" or base system as described in section 3.6. These real price results, along with the nominal price results to be discussed in section 4.10.2, are displayed graphically in Appendices E and F.

4.10.1 Observations

The first and most important observation is that the real prices of PCs have been dropping continuously between 1976 and 1987. In addition, the AAGR of -25.32% indicates that on average the quality-adjusted real prices have been dropping quite rapidly. In particular, real PC prices appear to be halving about every three years.

This does not necessarily mean, however, that the real prices of PCs have been decreasing this quickly. Since the above price index is quality-adjusted, it means that on the average, between adjacent years, prices drop by 25% while quality is unchanged, the quality attributes rise by 25% while price is unaffected, or some combination of the two occurs. This is remarkable and helps explain why market prices have not declined as fast as the index might indicate.

Another interesting point is that the quality-adjusted prices have not declined uniformly. Instead, the price declines have ranged from 2.67% in 1980 to 44.59% in 1977. Prices have been quite volatile, though always declining. It also appears that, aside from 1977, the price decline has been greater after 1982 than before it. This may be due to the effect of the IBM PC on the PC market since its introduction in late 1981. This possibility will be further examined in section 4.11.

Further, price declines appear to have been increasing recently, as can be seen in the price declines from 1985 to 1987. This might reflect dynamic changes in the PC market over this time, including the following: many new low-cost manufacturers, many new PC models, several important new developments such as the Apple Mac II and IBM PS/2, and the introduction of 32-bit PC architectures.

Since these trends seem to be continuing, if not accelerating, it might be reasonable to expect that the recent trend of high quality-adjusted real price decreases will continue into the foreseeable future.

4.10.2 Nominal Index

Using the real price index shown in Appendix D and the CPI data from section 3.4 it is possible to compute a nominal hedonic price index. This nominal index is worth reviewing since it allows a comparison between the results in this thesis and the results from earlier PC studies. In addition, since the nominal prices are the ones seen by PC purchasers, it will have more of an impact on their purchase behavior then will the real price index.

As shown in Appendices E and F, the real and nominal price indexes are very similar, although the real index declines consistently faster than the nominal index. In fact, in 1980 the nominal index rose by over 11%, while the real index declined slightly (-2.67%). This discrepancy is due to especially high inflation in 1979-1980 (14.2%).

The nominal PC price AAGR of -20.41% is noticeably less than the real price AAGR of -25.32%. This represents a halving of nominal PC prices every 3.7 years, slightly slower than is the case for the real prices.

Finally, the nominal price AAGR derived here is consistent with the results derived in Gordon (1987) and with some very early unpublished results by other researchers in this area.

4.11 Stratified Sample Set Analysis

In the previous section I examined the hedonic price index for all of the PCs within the data set. This price index is based on a "vanilla" PC system that has a 0 value for all of the dummy variables. Another way to examine the data is to stratify it into two or more groups and then run a regression on the data in each group to obtain a stratified hedonic price index. These stratified indexes provide more specific information about the price changes for the PCs within each data grouping.

In this section I report results of an analysis using three different criteria to divide the data set. First, by stratifying the data into 1976-1981 and 1982-1987 sets I shall see if IBM's late-1981 introduction of its PC had any effect on market prices.

Second, I stratified the data set into PCs sold at list price and those sold at a discount to determine whether the list prices or discount prices have been changing at different rates.

Finally, I stratified the data set by the size of the processor chip (8-bit vs. 16-bit vs. 32-bit) to see what effect the processor type has on prices.

For each of these stratifications, I display in Appendix D the values of the yearly dummy variables, the corresponding price index, and the resulting yearly price changes are displayed; for comparison, values based on the entire data set, reproduced in column one. Each of the stratified sample regression results is also examined to see how the results differ and what the differences might signify.

4.11.1 1976-1981 vs. 1982-1987

The first stratified sample run was set up by dividing the data into two groupings of six years each. This had several interesting effects on the resulting regressions and on the hedonic price index derived from the regressions.

First, the R^2 for the '76-'81 data was 88.8%, compared to 71.2% for the '82-'87 data. In addition, the SER (standard error of the regression) is smaller both in absolute terms and as a percentage of the mean of the dependent variable in 1976-1981, as can be seen in column two of Appendix D.

These results indicate that the regression model had greater predictive ability for the earlier time frame. This may be due to the fact that the PC market was much less diversified, fragmented, and competitive at that time. In addition, the average price decline during the later time segment is greater than in the earlier segment (28.97% vs. 26.73%), supporting the hypothesis that the PC prices have been declining at a faster rate in recent years.

In Appendix G I compare the original combined data set yearly price decreases with the results from these two stratified groups. During most of the years the price decreases are quite similar, except for 1979 and 1980, where the stratified data sample shows significantly smaller price declines. There is no obvious explanation for this difference, aside from the fact that the stratified data set is more specific but has fewer data points.

What is also interesting is that both the combined and stratified price indexes show a clear four year pricing wave with low points in 1980 and 1984, and some indications of an upcoming low point in 1988. This four year period may be related to the product life cycle for PCs, may just be an artifact of the data, or may represent the extremely dynamic nature of the PC market, which is still growing and changing.

In addition, the parameter values are clearly different in the stratified regressions, as shown in Appendices M and N. While a Chow test of parameter equality could not be run due to the differing independent variables in each regression,[33] many of the parameters (hard disk, slots, age) have very different coefficients

[33] Unfortunately, the modified Chow test used in sections 4.11.2 and 4.11.3 could also not be used, due to multicollinearity problems.

and t-statistics. This further reinforces the hypothesis that major changes have occurred in the market since the end of 1981.

4.11.2 <u>List Price vs. Discount Price</u>

Stratifying the data set by type of price also has significant effects on the regression and price index results. The R^2 for the discount price equation was noticeably greater than for the list price equation (82.1% vs. 71.0%). Also, as can be seen in column three of Appendix D, the SER is smaller for the discount prices, both in absolute and relative terms.

The stratified discount price data is better predicted by the model, perhaps because the discount price decreases have been much less than the list price decreases, at all times, though they have also been more volatile. This can be seen both in the yearly price index in Appendix I and in the yearly price decreases in Appendix H. The list price data closely follows the overall price data, while the discount prices are not only lower (as one would expect), but also decreasing at a lower rate.

The list prices dropped an average of 25.36%, while discount prices dropped only 17.83%. In addition, in two years (1980, 1984) discount prices actually rose. This is difficult to explain except that these years might have seen the removal of several deeply discounted PC models from the market, resulting in the average discount price increase.

These two years are also points of small price decreases across all of the data set stratifications and in the original combined data. Thus, some exogenous events (such as the Presidential elections) may have occurred that affected the PC market as a whole to some extent, and the discount PC market to a much greater extent.

In the last few years, however, discount prices have been dropping slightly faster than list prices have. This may well be due to the increasing competition in the PC market, resulting in greater and greater discounts from retailers and other middlemen.

Finally, the parameter values and t-statistics are quite different between the two stratified samples, as shown in Appendices O and P, though a Chow test could not be run to confirm this difference.[34] There are several anomalous results in the discount price regression which are particularly interesting. For example, 32-bit processors reduce a discount PC's price, color monitors have no impact, and the '79-'82 yearly dummy variables are not significant. With all of this, however, more of the variability is explained in the discount price regression than in the list price regression.

Thus, it appears that there are definite differences between the two data sets and that this is indeed a valid and interesting way to divide the PC data.

4.11.3 Processor Size (8, 16, 32 bit)

The final data set stratification is by the size of the microprocessor within the PC. As shown in column four of Appendix D, the R^2 and SER values are reasonably similar between the 16- and 32-bit systems, while rather different from the 8-bit systems. The 8-bit PC price index is very similar to the overall price index, although the other systems show lower price decreases for all years except 1984.

As can be seen in Appendix J, the 8-bit system prices have been dropping more quickly, particularly in recent years. This makes sense since 8-bit systems are now mostly obsolete and are being replaced by 16- and 32-bit systems. The 8-bit price

---

[34] A modified Chow test was run, however, using the two separate regressions as the unconstrained case and the original overall regression as the constrained case. An F-test was then done, using the test value determined by the error sum of squares of the regressions, as per a normal Chow test. Interestingly, the F-test value of .54 was less than the 5% confidence level of 1.47, indicating that it is not possible to rule out the possibility that the parameters are equal.

index is also at a much lower point, since the 8-bit prices have been dropping more quickly and because they have been on the market a much longer period of time.

The 16- and 32- bit systems have particularly low R^2 compared to the other stratified data sets, perhaps indicating that they are more varied and changing more quickly than the PCs in other stratifications. They are also newer, so they have only been available during the market upheavals of the last few years.

As can be seen in Appendices Q-S, the parameter values and t-statistics are quite different in the three data partitions, though a Chow test could not be run to confirm this observation.[35] In addition, several of the parameters in the 32-bit processor regression have the opposite sign from what is expected (RAM, slots, B&W, color), and most of the variables have low t-statistics and therefore appear insignificant.

The only possible explanation for this unusual result is that the 32-bit processor systems are so new and so different from other PCs that different variables must be used to analyze them. Also, it must be realized that there are comparatively few of these models on the market now, so some form of additional premium pricing may be occurring.

4.11.4 Conclusions

The above data stratifications provide further insights into the PC market and confirm the impression that quality-adjusted PC prices have been dropping rapidly within all segments. The hedonic price indexes generated for each of the stratified data sets provide unique insights, but their average price decreases are reasonably

---

[35] A modified Chow test, as described in section 4.11.2, was run, giving an F-test result of 4.01 which is greater than the 5% confidence level critical value of 1.47. This supports the hypothesis that the parameter coefficients are different between the different processor size regressions.

similar, as can be seen in Appendix L. It is readily apparent, however, that PC prices in '82-'87 have been dropping more rapidly and discount prices less rapidly than in the PC market as a whole.

One further stratification along the lines of the ones done above was attempted. The data set was stratified by the age of the system (0 or 1-7). Multicollinearity problems prevented a full analysis of the resulting regressions, but the partial results obtained suggested that the prices of new models dropped faster than the prices of models that had been on the market for at least one year (-24.7% AAGR for new models, -22.6% AAGR for "old" models). This is consistent with the parameter coefficient obtained for the Age variable, discussed in section 4.8.

It might also be illuminating to stratify the data on other variables such as whether the system has a hard disk and whether it is portable. The PC market data is rich enough for much further analysis of this type.

4.12 Residual Analysis

The final piece of analysis I undertook involved an examination of the residuals from the overall regression equation to see if the prices of any models were consistently over- or under-estimated. The PCs with the ten greatest residuals were as follows:

Model	Year	Residual
Sinclair 1500	1983	-2.01
Atari 800XL	1985	-1.87
Atari 800XL	1985	-1.84
Sinclair 1000	1983	-1.69
Atari 800XL	1986	-1.59
Sinclair 1000	1982	-1.48
Sinclair 1000	1983	-1.45
Sinclair 1000	1983	-1.31
PC Network PC+	1986	-1.27
TI 99-4	1983	-1.26

Several interesting items can be noted from the above list. First, all of the residuals are more than two standard deviations from the mean (SER = .43), and most are at least three standard deviations from the mean. This indicates that these residuals are significant.

Second, the PCs with the eight highest residuals are all from either Sinclair or Atari. It appears that these manufacturers may price lower than average and therefore a separate manufacturer dummy variable may be appropriate for them.

Third, all of the highest residuals are negative, indicating that the actual real prices for all of the above systems are less than what was estimated by the regression. It appears that manufacturers are more likely to discount deeply their systems than they are to put exceptionally heavy premiums on them, no doubt due to the very competitive nature of the PC market.

Finally, all of the above systems are at the low end of the PC spectrum, where the most competition and price cutting has occurred. Thus, it is not surprising that the regression model was less adept at predicting the prices of these models, since the exogenous effect of additional competition has had a proportionally greater effect on these models.

The information provided by the residuals gives some further insights into the PC market, particularly the low end. While this information could be helpful to someone purchasing PCs, there is a caveat that the larger residuals might well be due to one or more important omitted variables, incorrect specification of the functional form of the regression equation, time lag effects, or certain marketing practices.[36]

---

[36] Gordon, p. 19

5.0 Possible Extensions

Many possible extensions presented themselves as I compiled this study. One interesting extension would be to somehow factor in such intangible characteristics as maintenance, support and training into the regression equation. Personal computers are also often sold with a limited warranty, usually lasting 90 days to a year. These factors no doubt affect the price that is set by the manufacturer and that the customer is willing to pay, yet these factors have not been examined.

Such factors would need to be quantified in some fashion, evaluated for each model/time data point and then placed into the equation. A problem with doing this, however, is one we have previously seen; obtaining this data would be difficult to impossible for model/time data points older than a year or two. This type of data would almost certainly need to be obtained on a yearly basis, permitting future studies to use these possibly very important independent variables.

Yet another intangible factor that would likely affect computer prices is the quality, availability and ease of use of the software for each model. It is commonly agreed that the quality of software has improved greatly,[37] yet this is not currently taken into account within my hedonic price index. Software could be a very important factor, however, since "market wisdom" often says that a computer will rise or fall based on the software that is available for it. Indeed, one of the main strengths of the IBM PC is reputed to be the large base of software for it, while several other models (most notably the Apple Lisa) are thought to have been unsuccessful for exactly the opposite reason.

It would also be interesting to do more research to obtain additional data in the 1976-1979 time frame. The current data set is biased towards the 1984-1987 time

---

[37] Chow, p. 121.

frame due to both a proliferation of PCs at this time as well as it being easier to research the more recent PCs.

Another possible extension would involve several additional data stratifications, as suggested in section 4.11.4. This would provide further information on various sub-markets or niches within the overall PC market.

It would also be interesting to research the "speed variables" to find a common way of representing them within the regression equation. While often the MIPS (millions of instructions per second) value is used to compare the power of mainframe computers, no one indicator is quite as applicable in the PC industry, although the Norton Index has been used by some in this capacity. Variables such as MHZ, processor type, bus size, bus speed, RAM access speed, hard disk seek time, hard disk access time, floppy disk seek time, floppy disk access time, etc. are used to judge the relative power or throughput of PCs. It is usually difficult, however, to find these data for older PC models, and they are frequently measured on differing scales.

Optimally, running benchmarks on all of the PCs in the regression data base or using benchmarks that have been published by independent groups (such as PC Magazine) would be appropriate. This type of information is often difficult to obtain, however, given that there are so many PCs on the market and that only the most "significant" of these appear to be publicly benchmarked.

Finally, it would be quite interesting to attempt to forecast what will happen to the quality-adjusted prices of PCs over the next several years. This would not be a particularly easy task, however, due to the apparent non-equilibrium state of the PC market.

Thus, the personal computer market is a rich field for further study using regression analysis to form hedonic price indexes.

6.0 Conclusions

As we have seen, the quality-adjusted real prices for personal computers have been dropping at an annualized cumulative rate of about 25%. This is a reasonably high price decline, indicating that systems will halve in quality-adjusted real price approximately every three years.

Thus, in the eleven years that PCs have been on the market, prices have declined to about one-twenty-fifth of where they were initially. In addition, not only is this trend continuing, but the recent advent of new manufacturers, new models, new technologies, and greater competition appears to have accelerated the price decline. This is vividly seen in Appendix G.

I expect these price declines will continue, as do many other observers, who urger users to take them into account when purchasing PCs. Every two or three years a new "generation" of PCs has been appearing, resulting in a greater than average price decline. This occurred in 1984 and 1986 and will likely occur again in 1988 or 1989.

Thus, purchasers of PCs might do well to time their purchases just after a new advanced PC model becomes available, when the prices of older models often decline to a greater degree than normal. Many purchasers did just this when they ignored the IBM PS/2 models in 1986 and instead bought the older AT and AT-compatible models that were sold at much lower prices.

Manufacturers in the PC market should also be well aware of the trends in quality-adjusted PC pricing. Unless they are serving a niche segment such as portable PCs (Toshiba), very high end (Compaq), very low end (Sinclair), or something similar, they must be prepared to face rapid price declines and/or rapid quality enhancements of their PCs.

This implies that manufacturers must either be low-cost producers, have an excellent R&D department or else be very good at copying new enhancements introduced by other companies. Even though the PC market is growing rapidly, the amount of competition is growing even faster, and any manufacturer that does not have a significant competitive advantage is likely to do rather poorly.

The PC market can still be profitable, but the increasing level of competition and fast-decreasing quality-adjusted prices indicate that the market may not be able to support nearly as many manufacturers as currently exist, once the growth phase for the product begins to taper off.

The Hedonic Price Index is a powerful technique for examining the pricing policies over time of different PC models. It is more flexible than the "matched-model" technique and allows for many more data points to be used. The Hedonic Price Indexes computed in this paper have answered interesting issues concerning PC pricing, and have also provided insights on PCs in general.

As I indicated in section 5.0, there is still much more work to be done in this area, but hopefully this study has provided some illumination and seeds for further analysis.

7.0 Bibliography

Bell, Jack, "For What It's Worth," Personal Computing, March, 1988, p. 42.

Berndt, Ernst, "The Measurement of Quality Change: Constructing an Hedonic Price Index for Computers," Chapter 4 of **The Practice of Econometrics**, Addison-Wesley Publishing Company, Reading, MA, forthcoming.

Cale, E.G., L.L. Gremillions and J.L. McKenney, "Price/Performance Patterns of U.S. Computer Systems," Communications of the ACM, April, 1979, volume 22, number 4, pp. 225-233.

Cartwright, David, "Improved Deflation of Purchases of Computers," Survey of Current Business, March, 1986, pp. 7-9.

Chow, Gregory, "Technological Change and the Demand for Computers," American Economics Review, vol. 57, December, 1967, pp. 117-130.

Cole, Roseanne, Yung Chen, Joan Barquin-Stolleman, Ellen Dulberger, Nurhan Helvacian, and James Hodge, "Quality-Adjusted Price Indexes for Computer Processors and Selected Peripheral Equipment," Survey of Current Business, January, 1986, pp. 41-50.

Dulberger, Ellen, "The Application of an Hedonic Model to a Quality Adjusted Price Index for Computer Processors," unpublished Ph.D. dissertation, The City University of New York, 1986.

Gordon, Robert, "The Postwar Evolution of Computer Prices," Working Paper 227, National Bureau of Economic Research, Inc., April, 1987.

Pindyck, Robert, Daniel Rubinfeld, **Econometric Models & Economic Forecasts**, McGraw-Hill Book Company, New York, 1981.

Triplett, Jack, "Price and Technological Change in a Capital Good: A Survey of Research on Computers," Discussion Paper 23, Bureau of Economic Analysis, U.S. Department of Commerce, June, 1987.

8.0 Appendices

8.1 List of Appendices

8.2 Regression Variables

Variable	Full Name	Variable Number(Sect. 3.5)
LRPRIC	log of real price	-
C	constant	-
RAM	random access memory	1
MHZ	clock speed	4
HARDDI	hard disk size	5
NUMFLO	floppy drives	8
SLOTS	expansion slots	9,10,11
AGE	age of system	14
DPROC1	16-bit processor	15
DPROC3	32-bit processor	16
DBW	B&W monitor	17
DCOLOR	color monitor	18
DPORT	portability	19
DEXTRA	extra equipment	20
DDISCO	discount price	21
DAPPLE	Apple computer	22
DCOMMO	Commodore computer	23
DCOMPA	Compaq computer	24
DIBM	IBM computer	25
DNEC	NEC computer	26
DPCLIM	PC Limited computer	27
DRADIO	Radio Shack computer	28
D77	1977 data point	29
D78	1978 data point	30
D79	1979 data point	31
D80	1980 data point	32
D81	1981 data point	33
D82	1982 data point	34
D83	1983 data point	35
D84	1984 data point	36
D85	1985 data point	37
D86	1986 data point	38
D87	1987 data point	39

Appendix A - Initial Regression Results

LS // Dependent Variable is LRPRIC
Date: 3-08-1988 / Time: 9:58
SMPL range: 1 - 962
Number of observations: 960

VARIABLE	COEFFICIENT	STD. ERROR	T-STAT.	2-TAIL SIG.
C	6.2658311	0.1437409	43.591142	0.000
RAM	0.3720186	0.0238365	15.607117	0.000
MHZ	0.1438956	0.0419786	3.4278292	0.001
HARDDI	0.1831977	0.0120570	15.194272	0.000
NUMFLO	0.4223733	0.0510901	8.2672215	0.000
SLOTS	0.1974297	0.0210095	9.3971486	0.000
AGE	0.0969816	0.0329201	2.9459682	0.003
DPROC1	0.2258055	0.0408410	5.5288880	0.000
DPROC3	0.4179024	0.0868444	4.8120843	0.000
DBW	0.1493298	0.0321334	4.6471795	0.000
DCOLOR	0.1970422	0.1053714	1.8699779	0.061
DPORT	0.2525675	0.0466279	5.4166643	0.000
DEXTRA	0.2787576	0.0965089	2.8884139	0.004
DDISCO	-0.2942345	0.0348569	-8.4412065	0.000
DAPPLE	0.1123856	0.0668784	1.6804485	0.093
DCOMMO	-0.4405415	0.0739208	-5.9596407	0.000
DCOMPA	0.2758083	0.0644499	4.2794229	0.000
DIBM	-0.0717625	0.0526478	-1.3630668	0.173
DNEC	0.0026506	0.0806155	0.0328795	0.974
DPCLIM	-0.5533108	0.0968099	-5.7154333	0.000
DRADIO	-0.0482654	0.0563547	-0.8564564	0.392
D77	-0.5368305	0.1733804	-3.0962580	0.002
D78	-0.8007641	0.1772153	-4.5185951	0.000
D79	-0.9341432	0.1650566	-5.6595317	0.000
D80	-0.9468389	0.1591875	-5.9479495	0.000
D81	-1.0907076	0.1648055	-6.6181501	0.000
D82	-1.3972945	0.1631208	-8.5660118	0.000
D83	-1.8775559	0.1608053	-11.675959	0.000
D84	-1.9702135	0.1634130	-12.056650	0.000
D85	-2.4096350	0.1699993	-14.174380	0.000
D86	-2.8155441	0.1729214	-16.282220	0.000
D87	-3.1130191	0.1754684	-17.741196	0.000

R-squared	0.756872	Mean of dependent var	6.965944
Adjusted R-squared	0.748750	S.D. of dependent var	0.844155
S.E. of regression	0.423131	Sum of squared resid	166.1492
F-statistic	93.19083	Log likelihood	-520.2384

Appendix 3 - Final Regression Results

LS // Dependent Variable is LRPRIC
Date: 5-02-1988 / Time: 22:58
SMPL range: 1 - 1055
Number of observations: 1053

VARIABLE	COEFFICIENT	STD. ERROR	T-STAT.	2-TAIL SIG.
C	6.2775309	0.1449213	43.316813	0.000
RAM	0.3719182	0.0227098	16.376980	0.000
MHZ	0.2011689	0.0389994	5.1582511	0.000
HARDDI	0.1779840	0.0114676	15.520662	0.000
NUMFLO	0.4428315	0.0499662	8.8626161	0.000
SLOTS	0.1864704	0.0202784	9.1955378	0.000
AGE	0.0962747	0.0323227	2.9785455	0.003
DPROC1	0.1912688	0.0372574	5.1337134	0.000
DPROC3	0.3849790	0.0813086	4.7347859	0.000
DBW	0.1382360	0.0303074	4.5611354	0.000
DCOLOR	0.2028183	0.0913458	2.2203349	0.026
DPORT	0.2629334	0.0439145	5.9873886	0.000
DEXTRA	0.2941825	0.0916609	3.2094667	0.001
DDISCO	-0.2987001	0.0338351	-8.8281062	0.000
DAPPLE	0.1592269	0.0648456	2.4554761	0.014
DCOMMO	-0.3827383	0.0710642	-5.3858098	0.000
DCOMPA	0.2819372	0.0633874	4.4478410	0.000
DPCLIM	-0.5474602	0.0976176	-5.6082144	0.000
D77	-0.5903823	0.1734329	-3.4040962	0.001
D78	-0.8444652	0.1777240	-4.7515543	0.000
D79	-0.9798603	0.1635325	-5.9918391	0.000
D80	-1.0069511	0.1565563	-6.4318801	0.000
D81	-1.1760740	0.1618545	-7.2662445	0.000
D82	-1.4932413	0.1603157	-9.3143798	0.000
D83	-1.9858852	0.1576848	-12.594016	0.000
D84	-2.0533304	0.1600245	-12.831348	0.000
D85	-2.4839775	0.1662313	-14.942897	0.000
D86	-2.9051528	0.1684487	-17.246518	0.000
D87	-3.2108891	0.1707531	-18.804282	0.000

R-squared	0.745619	Mean of dependent var	6.973420
Adjusted R-squared	0.738663	S.D. of dependent var	0.840569
S.E. of regression	0.429709	Sum of squared resid	189.0810
F-statistic	107.1948	Log likelihood	-590.0243

Appendix C - Weighted Regression Results

LS // Dependent Variable is LRPRIC
Date: 3-24-1988 / Time: 11:57
SMPL range: 1 - 1020
Number of observations: 1018
Weighting series: FITT

VARIABLE	COEFFICIENT	STD. ERROR	T-STAT.	2-TAIL SIG.
C	6.2159209	0.1497473	41.509414	0.000
RAM	0.3913704	0.0227197	17.226066	0.000
MHZ	0.1675843	0.0391168	4.2842006	0.000
HARDDI	0.1818077	0.0121461	14.968412	0.000
NUMFLO	0.4411658	0.0506861	8.7038807	0.000
SLOTS	0.1979622	0.0205020	9.6557354	0.000
AGE	0.0924031	0.0324363	2.8487570	0.004
DPROC1	0.1936466	0.0387278	5.0001985	0.000
DPROC3	0.3612284	0.0877130	4.1182984	0.000
DBW	0.1710623	0.0307949	5.5548988	0.000
DCOLOR	0.2251513	0.0967720	2.3266173	0.020
DPORT	0.2658440	0.0443732	5.9910951	0.000
DEXTRA	0.2888522	0.0959441	3.0106287	0.003
DDISCO	-0.2914381	0.0341537	-8.5331367	0.000
DAPPLE	0.1609473	0.0656668	2.4509689	0.014
DCOMMO	-0.3806178	0.0665544	-5.7189010	0.000
DCOMPA	0.2913771	0.0668250	4.3603014	0.000
DPCLIM	-0.5542259	0.0987839	-5.6104851	0.000
D77	-0.5734665	0.1784780	-3.2130933	0.001
D78	-0.8229407	0.1818580	-4.5251821	0.000
D79	-0.9842555	0.1684604	-5.8426503	0.000
D80	-1.0076852	0.1613985	-6.2434624	0.000
D81	-1.1826907	0.1659793	-7.1255319	0.000
D82	-1.5139500	0.1646929	-9.1925630	0.000
D83	-2.0354403	0.1621758	-12.550824	0.000
D84	-2.0676223	0.1646379	-12.558604	0.000
D85	-2.5167230	0.1708189	-14.733282	0.000
D86	-2.9393824	0.1728346	-17.006907	0.000
D87	-3.2366718	0.1747712	-18.519480	0.000

Weighted Statistics

R-squared	0.439357	Mean of dependent var	6.925777
Adjusted R-squared	0.423485	S.D. of dependent var	0.573059
S.E. of regression	0.435116	Sum of squared resid	187.2433
F-statistic	27.68027	Log likelihood	-582.6475

Unweighted Statistics

R-squared	0.753534	Mean of dependent var	6.967340
Adjusted R-squared	0.746556	S.D. of dependent var	0.845219
S.E. of regression	0.425510	Sum of squared resid	179.0675

Appendix D - Hedonic Price Indexes

Year Parameter Values

	Strat. - Year			Strat. - Price		Strat. - Processor		
	76-87	76-81	82-87	Discount	List	8-bit	16-bit	32-bit
1976								
1977	-0.590	-0.646			-0.556	-0.553		
1978	-0.844	-0.900			-0.826	-0.826		
1979	-0.980	-1.192		-0.065	-0.974	-1.043		
1980	-1.007	-1.360		0.011	-1.036	-1.105		
1981	-1.176	-1.555		-0.139	-1.213	-1.259		
1982	-1.493			-0.376	-1.562	-1.642		
1983	-1.986		-0.460	-1.149	-1.903	-2.247	-0.267	
1984	-2.053		-0.513	-0.849	-2.086	-2.169	-0.617	
1985	-2.484		-0.951	-1.344	-2.504	-2.646	-0.925	
1986	-2.905		-1.371	-1.764	-2.917	-3.122	-1.300	
1987	-3.211		-1.710	-2.160	-3.217	-3.497	-1.590	-0.299
R-sq	74.60%	88.80%	71.80%	82.10%	71.00%	75.40%	64.40%	60.00%
SER	0.43	0.292	0.444	0.394	0.428	0.459	0.344	0.286
dep.	6.9734	6.8830	7.0047	6.6721	7.0886	6.6570	7.2260	7.9597
S/d	6.17%	4.24%	6.34%	5.91%	6.04%	6.89%	4.76%	3.59%

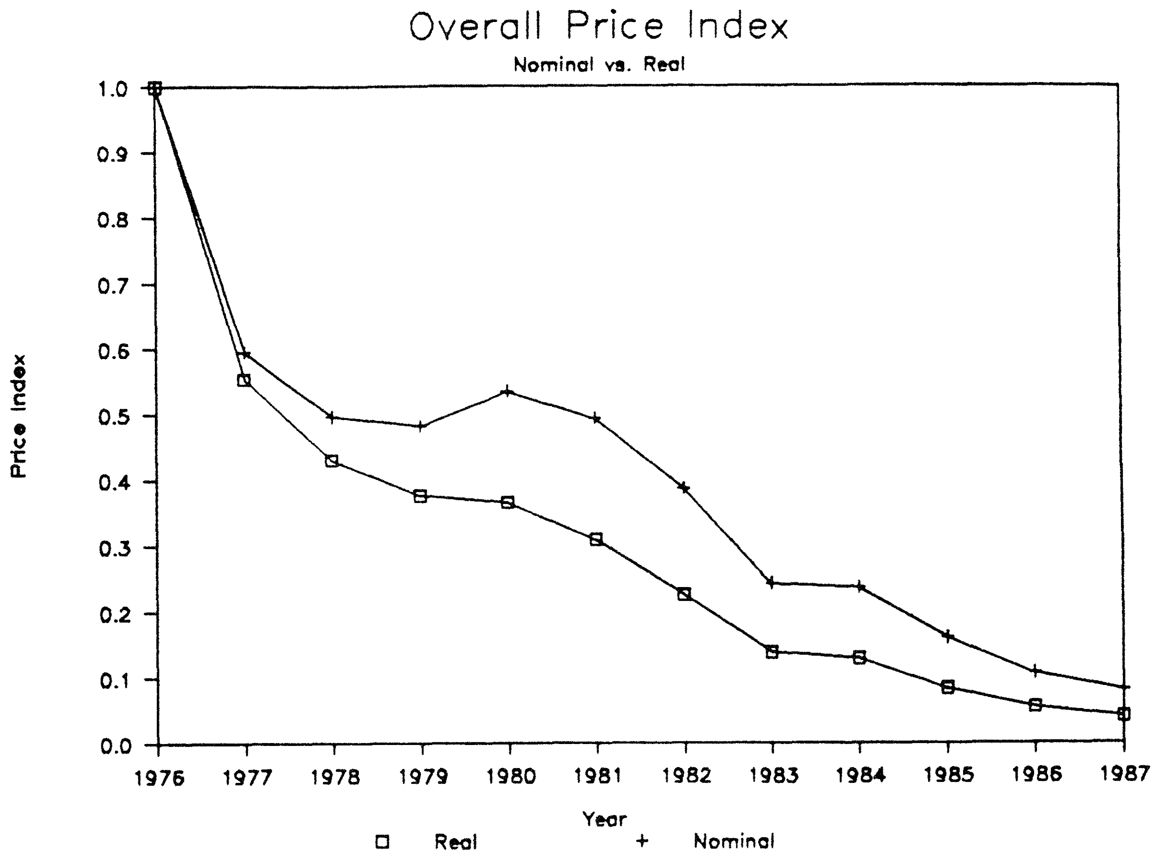
Price Indexes

	76-87	76-81	82-87	Discount	List	8-bit	16-bit	32-bit
1976	1.0000	1.0000			1.0000	1.0000		
1977	0.5541	0.5242			0.5732	0.5753		
1978	0.4298	0.4067		1.0000	0.4376	0.4376		
1979	0.3754	0.3036		0.9368	0.3776	0.3525		
1980	0.3653	0.2568		1.0115	0.3547	0.3311		
1981	0.3085	0.2112		0.8703	0.2974	0.2838		
1982	0.2246		1.0000	0.6869	0.2096	0.1936	1.0000	
1983	0.1373		0.6315	0.3169	0.1492	0.1057	0.7658	
1984	0.1283		0.5989	0.4276	0.1242	0.1143	0.5396	
1985	0.0834		0.3863	0.2609	0.0817	0.0710	0.3964	
1986	0.0547		0.2539	0.1713	0.0541	0.0441	0.2726	1.0000
1987	0.0403		0.1808	0.1154	0.0401	0.0303	0.2040	0.7412

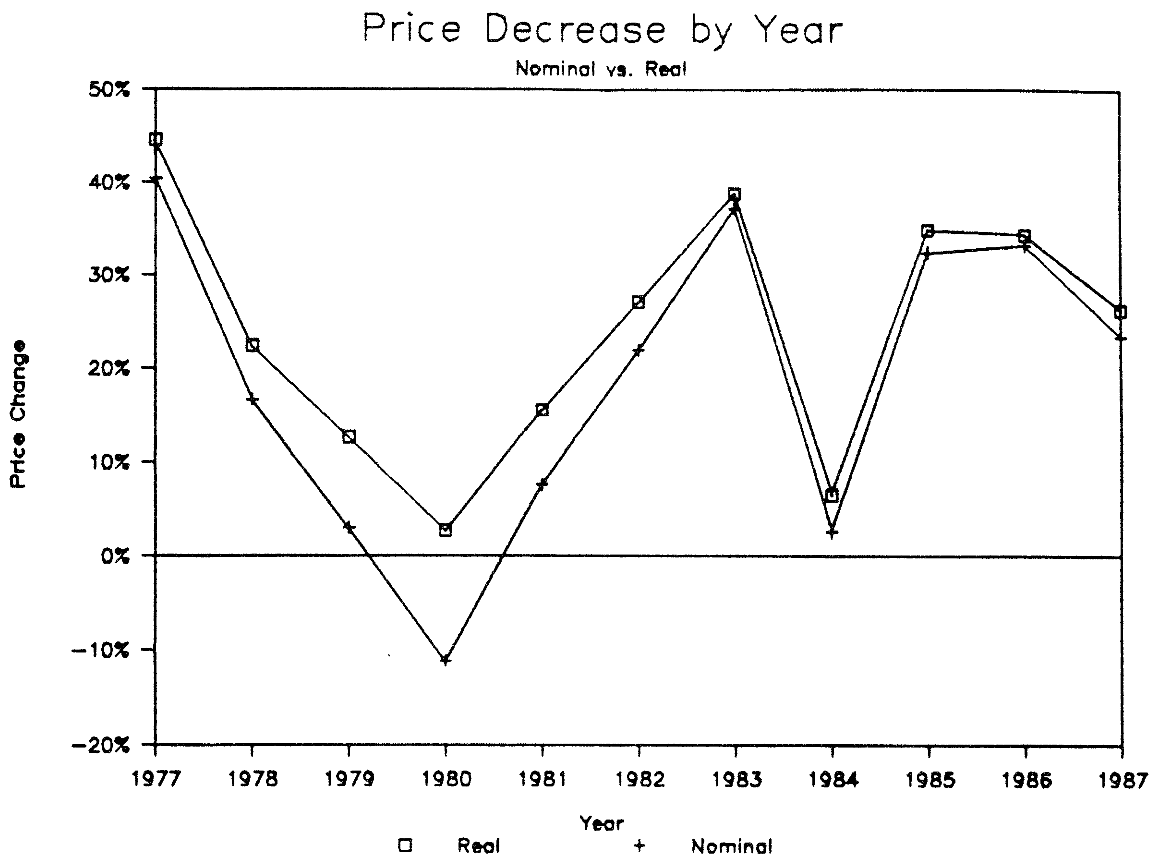
Price Decline

	76-87	76-81	82-87	Discount	List	8-bit	16-bit	32-bit
1976								
1977	44.59%	47.58%			42.68%	42.47%		
1978	22.44%	22.42%			23.66%	23.93%		
1979	12.66%	25.36%		6.32%	13.72%	19.44%		
1980	2.67%	15.41%		-7.97%	6.05%	6.07%		
1981	15.56%	17.75%		13.96%	16.16%	14.28%		
1982	27.18%			21.07%	29.51%	31.80%		
1983	38.90%		36.85%	53.86%	28.84%	45.41%	23.42%	
1984	6.52%		5.15%	-34.93%	16.76%	-8.17%	29.54%	
1985	34.99%		35.50%	38.99%	34.17%	37.92%	26.53%	
1986	34.37%		34.27%	34.33%	33.85%	37.87%	31.22%	
1987	26.34%		28.78%	32.66%	25.92%	31.31%	25.17%	25.88%
Cum.	25.32%	26.73%	28.97%	17.83%	25.36%	27.23%	27.23%	25.88%

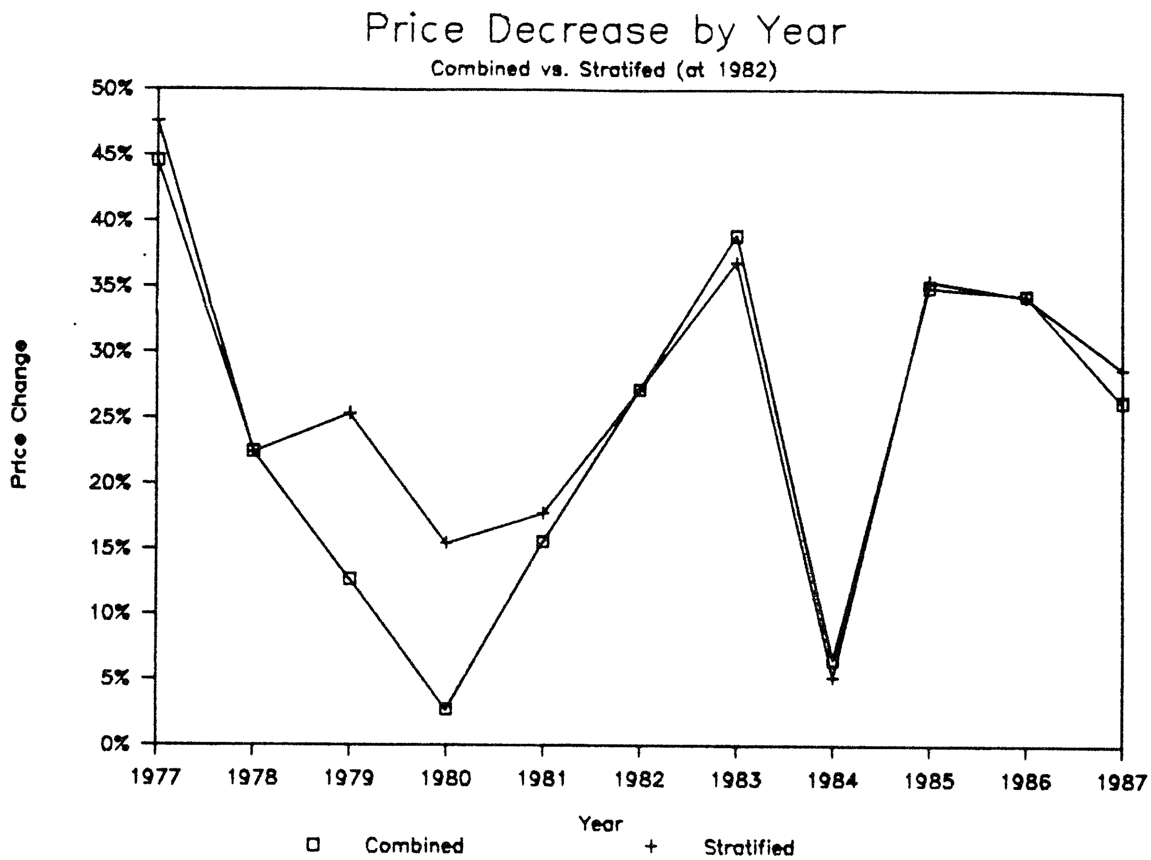
Appendix E


Overall Price Index
Nominal vs. Real
Price Index
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
Year
Real
Nominal

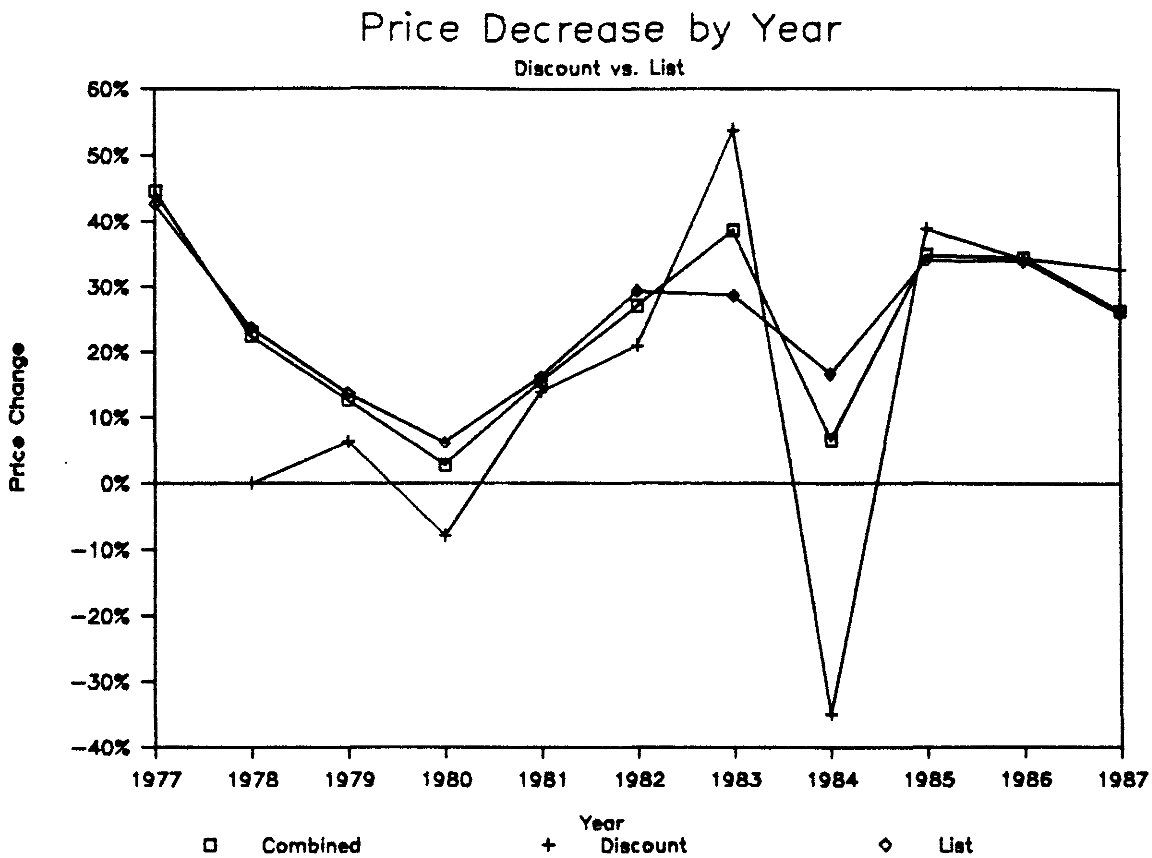
Appendix F


Price Decrease by Year
Nominal vs. Real
50%
40%
30%
20%
10%
0%
-10%
-20%
Price Change
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
Year
Real
Nominal

Appendix G

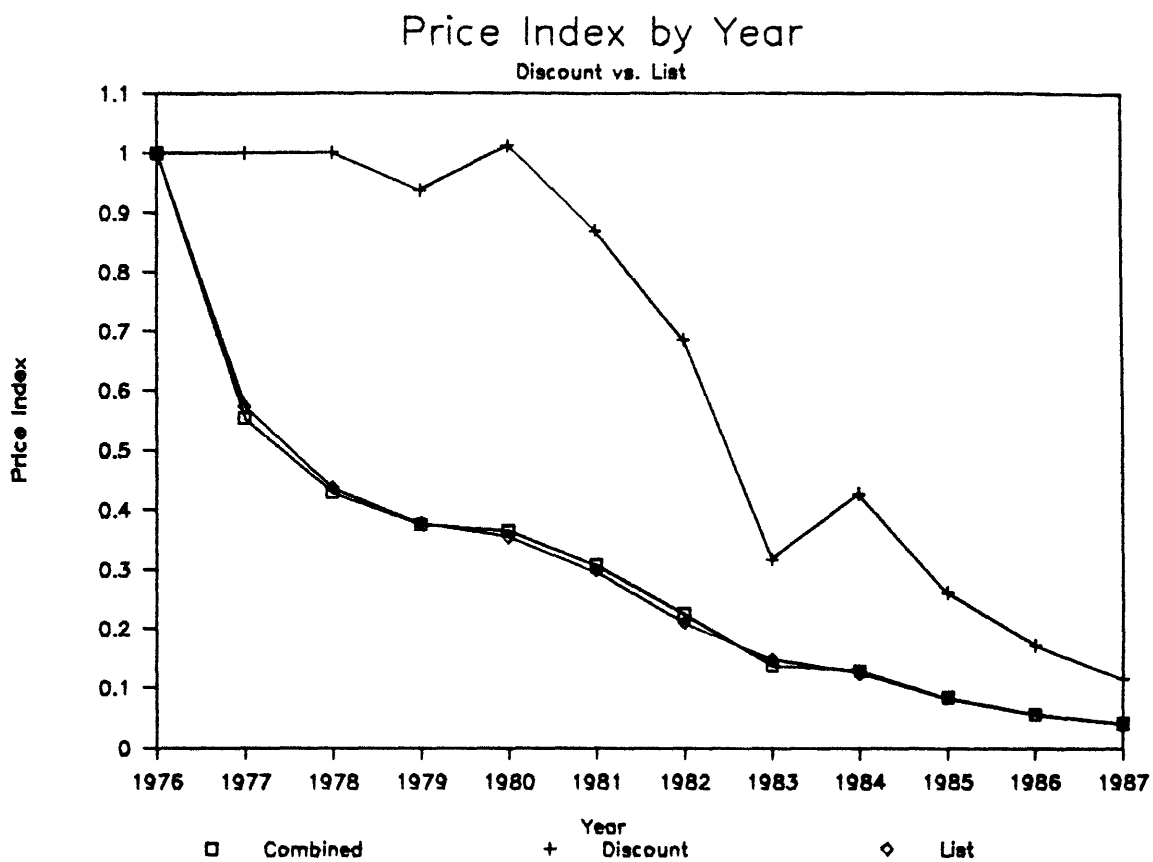

Price Decrease by Year
Combined vs. Stratifed (at 1982)
50%
45%
40%
35%
30%
25%
20%
15%
10%
5%
0%
Price Change
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
Year
Combined
Stratified

Appendix H



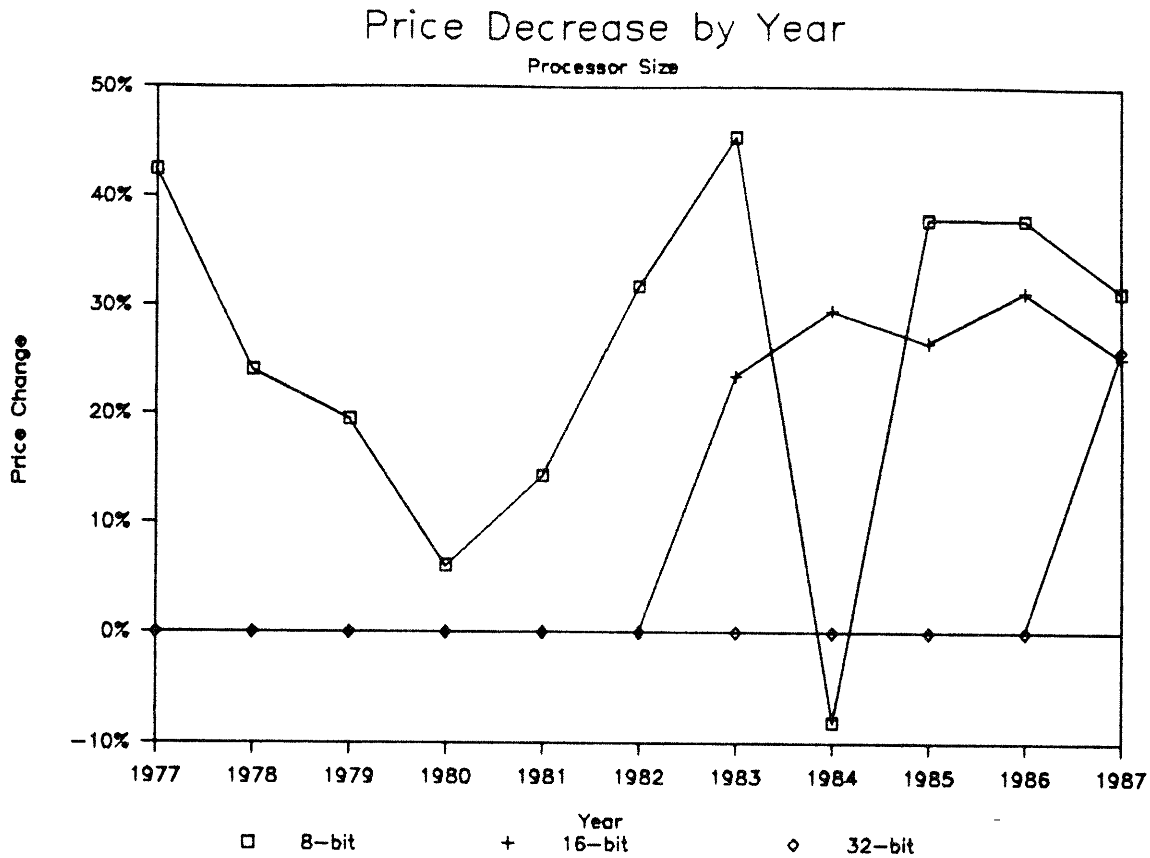
Price Decrease by Year
Discount vs. List
60%
50%
40%
30%
20%
10%
0%
-10%
-20%
-30%
-40%
Price Change
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
Year
Combined
Discount
List

Appendix I



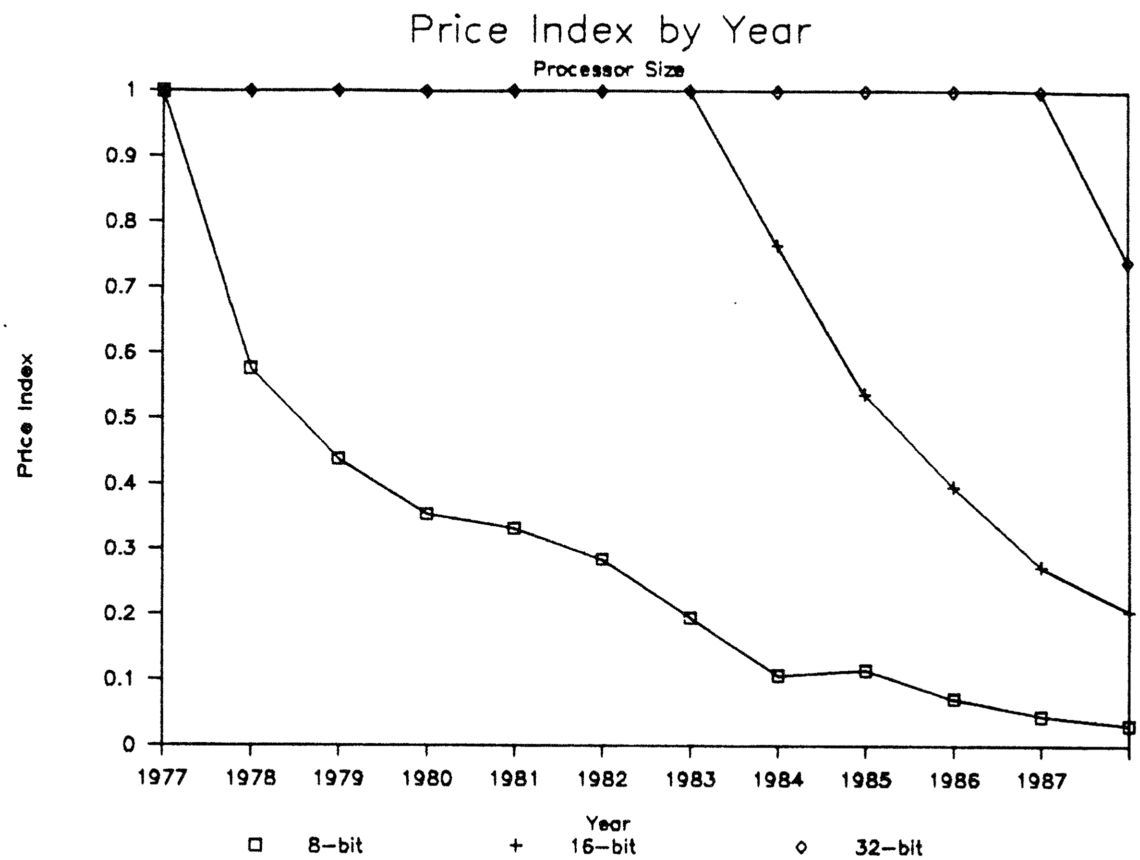
Price Index by Year
Discount vs. List
Price Index
1.1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
Year
Combined
Discount
List

Appendix J

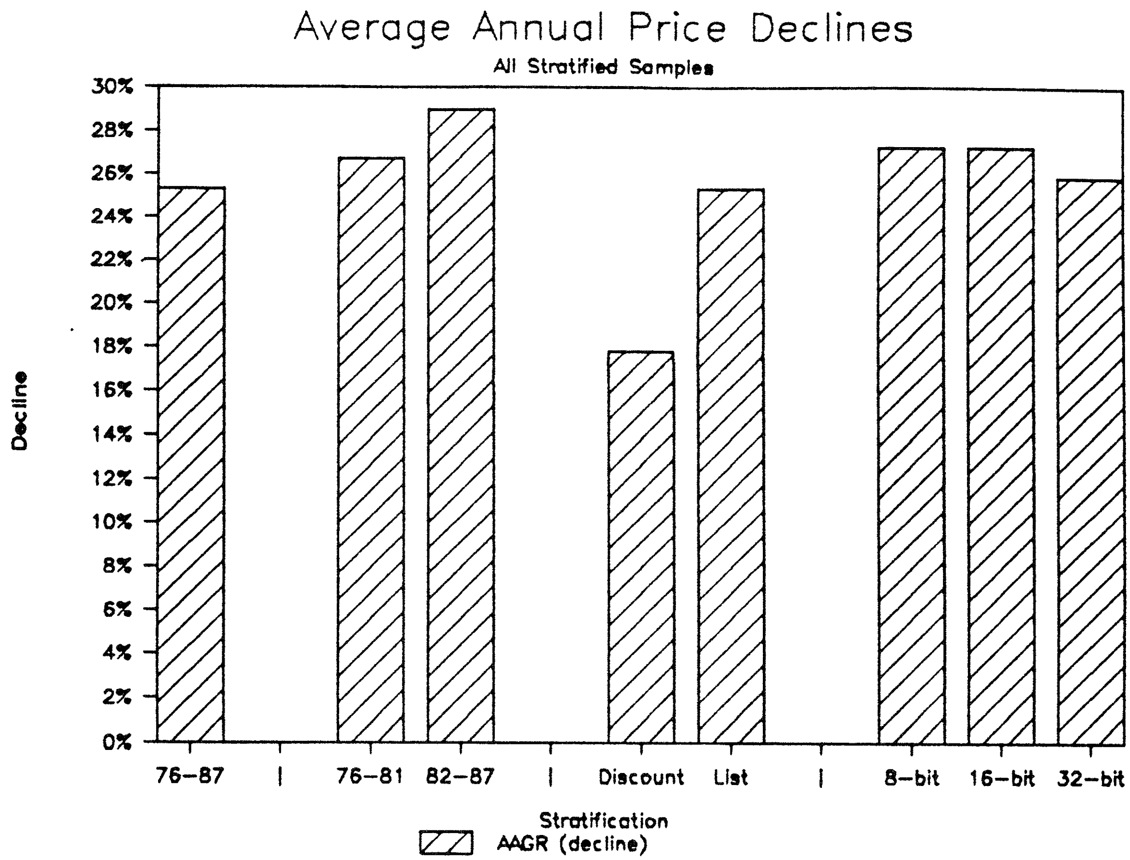


Price Decrease by Year
Processor Size
50%
40%
30%
20%
10%
0%
−10%
Price Change
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
Year
8−bit
16−bit
32−bit

Appendix K


Price Index by Year
Processor Size
Price Index
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
Year
8–bit
16–bit
32–bit

Appendix L


Average Annual Price Declines
All Stratified Samples
30%
28%
26%
24%
22%
20%
18%
16%
14%
12%
10%
8%
6%
4%
2%
0%
Decline
76-87
76-81
82-87
Discount
List
8-bit
16-bit
32-bit
Stratification
AAGR (decline)

Appendix H - Stratified ('76-'81) Regression Results

LS // Dependent Variable is LRPRIC
Date: 3-25-1988 / Time: 16:50
SMPL range: 1 - 156
Number of observations: 155

VARIABLE	COEFFICIENT	STD. ERROR	T-STAT.	2-TAIL SIG.
C	6.6639347	0.1306545	51.004250	0.000
RAM	0.3561644	0.0347901	10.237525	0.000
MHZ	0.2556632	0.0711372	3.5939464	0.000
HARDDI	0.7436057	0.0654449	11.362325	0.000
NUMFLO	0.8909930	0.0856133	10.407178	0.000
SLOTS	-0.0680375	0.0401505	-1.6945635	0.090
AGE	0.0732138	0.0695241	1.0530702	0.292
DDISCO	-0.2015523	0.0586975	-3.4337482	0.001
DBW	0.1450665	0.0641717	2.2606006	0.024
DCOLOR	0.4028950	0.1606656	2.5076612	0.012
DPORT	-0.4659303	0.1791539	-2.6007265	0.009
DEXTRA	-0.2203233	0.3220045	-0.6842244	0.494
DAPPLE	0.4902863	0.1240685	3.9517390	0.000
DCOMMO	0.0597502	0.1086696	0.5498339	0.582
D77	-0.6458202	0.1253415	-5.1524835	0.000
D78	-0.8997371	0.1291103	-6.9687472	0.000
D79	-1.1921861	0.1325164	-8.9965160	0.000
D80	-1.3595112	0.1373471	-9.8983596	0.000
D81	-1.5548767	0.1527056	-10.182188	0.000

R-squared	0.888300	Mean of dependent var	6.883041
Adjusted R-squared	0.873517	S.D. of dependent var	0.820470
S.E. of regression	0.291796	Sum of squared resid	11.57972
F-statistic	60.08617	Log likelihood	-18.88734

Appendix N - Stratified ('82-'87) Regression Results

LS // Dependent Variable is LRPRIC
Date: 3-25-1988 / Time: 16:57
SMPL range: 1 - 955
Number of observations: 951

VARIABLE	COEFFICIENT	STD. ERROR	T-STAT.	2-TAIL SIG.
C	4.6414132	0.1162903	39.912301	0.000
RAM	0.3571321	0.0263172	13.570289	0.000
MHZ	0.3553011	0.0423235	8.3948943	0.000
HARDDI	0.1879610	0.0112028	16.777974	0.000
NUMFLO	0.3729677	0.0562649	6.6287753	0.000
SLOTS	0.2408892	0.0231629	10.399778	0.000
AGE	0.1269601	0.0362873	3.4987453	0.000
DDISCO	-0.2674232	0.0369484	-7.2377390	0.000
DBW	0.0962737	0.0339995	2.8316249	0.005
DCOLOR	0.1967302	0.1027022	1.9155405	0.055
DPORT	0.3704326	0.0459172	8.0674018	0.000
DEXTRA	0.3205408	0.0966307	3.3171747	0.001
DAPPLE	0.2830838	0.0737609	3.8378566	0.000
DCOMMO	-0.3346671	0.0873747	-3.8302527	0.000
D83	-0.4597248	0.0811946	-5.6620091	0.000
D84	-0.5125798	0.0799169	-6.4139083	0.000
D85	-0.9510445	0.0855190	-11.120851	0.000
D86	-1.3706536	0.0874241	-15.678207	0.000
D87	-1.7101107	0.0884938	-19.324631	0.000

R-squared	0.718257	Mean of dependent var	7.004697
Adjusted R-squared	0.712816	S.D. of dependent var	0.828453
S.E. of regression	0.443964	Sum of squared resid	183.7012
F-statistic	131.9990	Log likelihood	-567.5917

Appendix O - Stratified (discount price) Regression Results

LS // Dependent Variable is LRPRIC
Date: 3-25-1988 / Time: 16:23
SMPL range: 1 - 268
Number of observations: 268

VARIABLE	COEFFICIENT	STD. ERROR	T-STAT.	2-TAIL SIG.
C	4.9038483	0.2599052	18.867835	0.000
RAM	0.3363525	0.0436799	7.7004000	0.000
MHZ	0.4059470	0.0743622	5.4590531	0.000
HARDDI	0.1887759	0.0268174	7.0393149	0.000
NUMFLO	0.2922862	0.0949302	3.0789599	0.002
SLOTS	0.2207480	0.0355042	6.2175207	0.000
AGE	0.0701093	0.0578220	1.2125018	0.225
DPROC1	0.0231161	0.0798725	0.2894123	0.772
DPROC3	-0.0601468	0.2333032	-0.2578053	0.797
DBW	0.1760078	0.0596466	2.9508453	0.003
DCOLOR	0.1931792	0.1619093	1.1931317	0.233
DPORT	0.2408505	0.0772139	3.1192645	0.002
DEXTRA	0.2144404	0.1930514	1.1107944	0.267
DAPPLE	0.3044015	0.1039305	2.9288943	0.003
DCOMMO	-0.3113848	0.1149693	-2.7084161	0.007
DCOMPA	0.1992023	0.0915102	2.1768325	0.029
D79	-0.0652720	0.2941481	-0.2219017	0.824
D80	0.0113926	0.2580457	0.0441494	0.965
D81	-0.1389678	0.2578941	-0.5388561	0.590
D82	-0.3756157	0.2622081	-1.4325100	0.152
D83	-1.1490987	0.2667511	-4.3077558	0.000
D84	-0.8494896	0.2658648	-3.1951944	0.001
D85	-1.3435752	0.2773278	-4.8447188	0.000
D86	-1.7641078	0.2831048	-6.2312885	0.000
D87	-2.1595626	0.2840165	-7.6036529	0.000

R-squared	0.821404	Mean of dependent var	6.672073
Adjusted R-squared	0.803765	S.D. of dependent var	0.889905
S.E. of regression	0.394214	Sum of squared resid	37.76336
Durbin-Watson stat	1.859025	F-statistic	46.56711
Log likelihood	-117.6827		

Appendix P - Stratified (list price) Regression Results

LS // Dependent Variable is LRPRIC
Date: 3-25-1988 / Time: 16:39
SMPL range: 1 - 840
Number of observations: 838

VARIABLE	COEFFICIENT	STD. ERROR	T-STAT.	2-TAIL SIG.
C	6.2946508	0.1473989	42.704867	0.000
RAM	0.3891406	0.0257772	15.096324	0.000
MHZ	0.1331596	0.0452764	2.9410386	0.003
HARDDI	0.1792045	0.0121342	14.768508	0.000
NUMFLO	0.4835503	0.0580048	8.3363881	0.000
SLOTS	0.1711741	0.0239354	7.1515149	0.000
AGE	0.1261785	0.0378221	3.3361094	0.001
DPROC1	0.2194312	0.0404387	5.4262714	0.000
DPROC3	0.4660953	0.0840528	5.5452709	0.000
DBW	0.0853871	0.0342888	2.4902336	0.013
DCOLOR	0.2122622	0.1076889	1.9710680	0.049
DPORT	0.2851088	0.0516262	5.5225561	0.000
DEXTRA	0.3255986	0.1032178	3.1544799	0.002
DAPPLE	0.1172615	0.0836604	1.4016361	0.161
DCOMMO	-0.3683601	0.0881719	-4.1777510	0.000
DCOMPA	0.3552433	0.0869845	4.0839844	0.000
DPCLIM	-0.5360745	0.0978538	-5.4783188	0.000
D77	-0.5564714	0.1736573	-3.2044222	0.001
D78	-0.8264680	0.1855289	-4.4546586	0.000
D79	-0.9739930	0.1690254	-5.7624060	0.000
D80	-1.0363545	0.1658430	-6.2490079	0.000
D81	-1.2126301	0.1754220	-6.9126434	0.000
D82	-1.5623386	0.1707690	-9.1488395	0.000
D83	-1.9025582	0.1634414	-11.640612	0.000
D84	-2.0860126	0.1666034	-12.520830	0.000
D85	-2.5041461	0.1742937	-14.367392	0.000
D86	-2.9174409	0.1768984	-16.492181	0.000
D87	-3.2174219	0.1798845	-17.886039	0.000

R-squared	0.709526	Mean of dependent var	7.088571
Adjusted R-squared	0.699843	S.D. of dependent var	0.781345
S.E. of regression	0.428072	Sum of squared resid	148.4288
F-statistic	73.27944	Log likelihood	-463.8180

Appendix Q - Stratified (8-bit proc.) Regression Results

LS // Dependent Variable is LRPRIC
Date: 5-02-1988 / Time: 23:09
SMPL range: 1 - 543
Number of observations: 542

VARIABLE	COEFFICIENT	STD. ERROR	T-STAT.	2-TAIL SIG.
C	6.1758229	0.1606033	38.453892	0.000
RAM	0.4465113	0.0307103	14.539440	0.000
MHZ	0.1332410	0.0581911	2.2897165	0.022
HARDDI	0.2748781	0.0267772	10.265379	0.000
NUMFLO	0.4461213	0.0675363	6.6056478	0.000
SLOTS	0.1388287	0.0313610	4.4267993	0.000
AGE	0.1216063	0.0465716	2.6111661	0.009
DBW	0.2438371	0.0441259	5.5259399	0.000
DCOLOR	0.4008665	0.1557213	2.5742567	0.010
DPORT	0.2434072	0.0605060	4.0228598	0.000
DEXTRA	0.2345956	0.1278368	1.8351177	0.066
DDISCO	-0.2854797	0.0467363	-6.1083113	0.000
DAPPLE	0.1752968	0.1144249	1.5319823	0.126
DCOMMO	-0.4031921	0.0890147	-4.5294999	0.000
DCOMPA	0.4220134	0.1408874	2.9953939	0.003
DPCLIM	-1.1235871	0.2135015	-5.2626664	0.000
D77	-0.5528703	0.1879431	-2.9416896	0.003
D78	-0.8263791	0.1920357	-4.3032578	0.000
D79	-1.0426055	0.1791973	-5.8181971	0.000
D80	-1.1052772	0.1745240	-6.3330974	0.000
D81	-1.2593742	0.1816769	-6.9319459	0.000
D82	-1.6421069	0.1842576	-8.9120150	0.000
D83	-2.2474511	0.1816657	-12.371355	0.000
D84	-2.1688839	0.1857285	-11.677712	0.000
D85	-2.6455754	0.1958271	-13.509748	0.000
D86	-3.1215854	0.2049383	-15.231830	0.000
D87	-3.4970919	0.2065934	-16.927415	0.000

R-squared	0.754349	Mean of dependent var	6.657007
Adjusted R-squared	0.741947	S.D. of dependent var	0.903769
S.E. of regression	0.459105	Sum of squared resid	108.5502
F-statistic	60.82575	Log likelihood	-333.2822

Appendix R - Stratified (16-bit proc.) Regression Results

LS // Dependent Variable is LRPRIC
Date: 5-02-1988 / Time: 23:16
SMPL range: 544 - 1047
Number of observations: 503

VARIABLE	COEFFICIENT	STD. ERROR	T-STAT.	2-TAIL SIG.
C	5.4841988	0.2302483	23.818629	0.000
RAM	0.2112701	0.0340063	6.2126701	0.000
MHZ	0.4144706	0.0645619	6.4197422	0.000
HARDDI	0.1602772	0.0111379	14.390202	0.000
NUMFLO	0.2729212	0.0781655	3.4915815	0.000
SLOTS	0.2486688	0.0302390	8.2234399	0.000
AGE	0.1344143	0.0427494	3.1442370	0.002
DBW	-0.0095472	0.0387837	-0.2461657	0.806
DCOLOR	0.2114444	0.1254874	1.6849851	0.092
DPORT	0.4011240	0.0596262	6.7273094	0.000
DEXTRA	0.3818129	0.1214746	3.1431504	0.002
DDISCO	-0.2829309	0.0429457	-6.5881076	0.000
DAPPLE	0.3983432	0.0845304	4.7124263	0.000
DCOMMO	-0.3090843	0.1466672	-2.1073852	0.035
DCOMPA	0.2847817	0.0669089	4.2562598	0.000
DPCLIM	-0.3620851	0.1013240	-3.5735389	0.000
D83	-0.2668867	0.1615222	-1.6523220	0.098
D84	-0.6170120	0.1544755	-3.9942373	0.000
D85	-0.9253063	0.1578407	-5.8622815	0.000
D86	-1.2996154	0.1582398	-8.2129485	0.000
D87	-1.5895845	0.1594585	-9.9686438	0.000

R-squared	0.643936	Mean of dependent var	7.226046
Adjusted R-squared	0.629162	S.D. of dependent var	0.566138
S.E. of regression	0.344758	Sum of squared resid	57.28970
F-statistic	43.58452	Log likelihood	-167.3500

Appendix S - Stratified (32-bit proc.) Regression Results

LS // Dependent Variable is LRPRIC
Date: 5-02-1988 / Time: 23:17
SMPL range: 1048 - 1108
Number of observations: 61

VARIABLE	COEFFICIENT	STD. ERROR	T-STAT.	2-TAIL SIG.
C	8.8413805	1.1356734	7.7851439	0.000
RAM	-0.1256735	0.1222188	-1.0282666	0.309
MHZ	0.0023976	0.2251863	0.0106473	0.992
HARDDI	0.1736799	0.0280697	6.1874418	0.000
NUMFLO	0.3758144	0.2568996	1.4628845	0.150
SLOTS	-0.2453749	0.1505324	-1.6300472	0.110
AGE	0.5387434	0.6279876	0.8578886	0.395
DBW	-0.3971501	0.1800953	-2.2052223	0.032
DCOLOR	-0.0457091	0.1339336	-0.3412816	0.734
DPORT	0.5449355	0.5804953	0.9387423	0.353
DDISCO	-0.4424896	0.1811789	-2.4422797	0.018
DAPPLE	0.2329501	0.2190288	1.0635595	0.293
DCOMPA	-0.1747357	0.4003385	-0.4364698	0.665
DPCLIM	0.0528326	0.2358627	0.2239973	0.824
D87	-0.2994301	0.3513857	-0.8521408	0.399

R-squared	0.600775	Mean of dependent var	7.959665
Adjusted R-squared	0.479272	S.D. of dependent var	0.395786
S.E. of regression	0.285605	Sum of squared resid	3.752232
Durbin-Watson stat	1.951383	F-statistic	4.944514
Log likelihood	-1.505301		