Updating the Hedonic Equations for the Price of Computers

By Fred Barzyk

Prices Division
Statistics Canada

Jean Talon Building, 13-D2, Ottawa, K1A 0T6

Facsimile: (613) 951-2848

Telephone: (613) 951-2493

Email: ***fred.barzyk@statcan.ca***

November 1999

Updating the Hedonic Equations for the Price Computers

1. Introduction[1]

The purpose of this paper is to give a brief description of the work that was recently carried out in Prices Division at Statistics Canada pertaining to the estimation of a new set of hedonic equations to be used in the treatment of quality changes for computers.[2] Generally speaking, hedonic equations attempt to explain the changes in a variable for a good or service by relating them to changes in the characteristics of the good or service. In this case, the relationship between the price of computers and the various characteristics of a computer, such as RAM, CPU speed, hard drive size and a number of qualitative variables (vendor, presence of a compact disc or CD drive, modem, etc.,) was measured. These estimations were carried out for desktop and notebook computers.[3]

The major problem pertaining to the estimation of these types of equations is the choice of functional form, since this choice will determine the quality of the results (relevance, accuracy, reliability, plausibility, etc.,). While almost all of the studies of hedonic equations for computers have limited the choice of functional form to considering either linear, semi-log or double log models, Triplett (1987) advocates the investigation of non-linear functional forms as well, since some of these have more economic meaning.

In the past, Prices Division has carried out three estimations of hedonic equations for computers (see Table 1). The first was carried out in 1989/90 and the second in 1991 and both used data from the Data Pro Reports, a database from the United States. The results of these studies favoured semi-log and linear functional forms for all of the desktop models estimated (16-bit, 32-bit and combined) respectively. The third set of equations was completed in 1996 and used data from the Public Works and Government Services Micro Acquisition Guide. This study resulted in the use of a semi-log model for quality adjustment to desktop computers, although no suitable model was found for notebook computers.

This latest revision of the hedonic equations for computers offers more relevant and interesting results than the previous two studies mentioned. More relevant because the data used is of much better quality than in the preceding analyses and more interesting in that a non-linear (in the parameters) model was tested in addition to linear, semi-log and double-log functions. The remainder of this paper is organised as follows: description of data, overview of models tested, methodology, results obtained and conclusion.

2. Data And Variables

The primary source of data for this study was the monthly pricing reports for computers supplied by International Data Corporation (IDC) of Canada to Prices Division. In comparison to the previous studies, the number of observations has greatly increased and the data is more relevant because it is the same data actually used produce the computer price index series.[4]

---

[1] *This paper should not be quoted without the explicit permission of the author. The author would like to express their gratitude to Andy Baldwin and Robin Lowe Prices Division as well as Erwin Diewert, Jack Triplett and Ralph Turvey for their helpful comments and suggestions. Any views expressed are those of the author and do not necessarily represent the opinions of Prices Division or Statistics Canada.*

[2] The focus of the paper is to describe the methodology and results obtained in the re-estimation of the equations and not to discuss at length the theory of hedonics. For a review of literature on the subject, the reader should consult *The Practice of Econometrics, Classic and Contemporary* by Ernst R. Berndt (1991), Chapter 4.

[3] The definition of a notebook computer is a computer weighing at least 2 pounds but less than 8 pounds (*Microsoft Press® Computer Dictionary,* pages 272-273).

[4] Given the limited data availability for the previous two studies, in the first case, American data had to be used while in the second study, the data came from the standing offers of sale to the Canadian Federal

A second source of data was the Internet site *'The CPU Scorecard'* (see Appendix A) which provides performance scores for CPUs based on the results of industry-wide benchmark performance tests. These scores were used to construct the variable SCORE, which is useful since it allows for quality changes to be carried out when models change generations of CPUs.[5]

2.1 - Desktop Computers

In the case of desktop computers, the data consisted of 576 observations. The variables found to be significant in explaining the differences in prices are presented in Table 2. Some of the variables are quantitative (e.g. the variable HD, which measures the size of the hard drive in megabytes), while others are dichotomous and therefore are treated as dummy variables (e.g. the variable MODEM, which represents the presence or absence of this feature). Finally, the variable BQUAL was constructed in accordance to industry quality reports rating the reliability and service of various computer companies.[6]

2. 2 - Notebook Computers

For the analysis of notebook computers, there were 174 observations for eight variables (see Table 3). Some were similar to those used in the analysis for desktops (RAM, HD, SCORE and XCACHE) while others pertained to notebooks only, namely WGHT, DSIZE and DQUAL. The variable HIEND separates the vendor IBM from the other vendors in the data set, reflecting the fact that IBM has traditionally offered notebook computers to the high-end user at higher prices than the competition.[7]

3. Methodology

In order to produce a new set of estimates for hedonic equations, two issues needed to be resolved, namely determining the explanatory variables and choosing an appropriate functional form. One is not independent of the other and they are often resolved in tandem.[8] In order to address the issue of functional form, it is important to realise that while there are several choices of functional form available, and for the most part they can be divided into two groups, linear and non-linear.[9] At the outset, the general model can be written as,

---

Government. However in this study, street prices for all observations were used, which are more reflective of actual market transaction prices.

[5] The inclusion of this variable allows us to estimate for the whole data set at once, rather than having to subdivide observations based on similar CPUs. In the past, model changes in the CPU were not reflected in the quality adjustment because they were not represented in any meaningful way. For example, the 1990 estimation separated the 16 and 32 –bit machines while the 1996 estimation was carried out for broad class of Pentium computers only.

[6] See *PC World, March 1996, pages 195-197*.

[7] This grouping of companies differs from that used in desktops, where companies were separated based on the perceived quality of their desktop product lines. In the market for notebooks, there is very little difference in the overall quality and service between the several notebook manufacturers contained in this analysis (see *PC World, November 1998,* pages 110 to 116).

[8] The contribution to the overall explanatory power of the independent variables was considered for *all* variables in *all* functional forms tested.

[9] There can be confusion when using the terms 'linear' and 'non-linear' with respect to the estimation of a functional form. Linear in this sense means linear in the parameters to be estimated. Logarithmic (or log-linear), reciprocal and exponential transformations are simple transformations to the general linear model. For example, a double-log function, while non-linear in its functional form, is still linear in its parameters and therefore can be estimated using ordinary least squares. Functional forms that are non-linear in their parameters though, as in the "t-identification" model, require some other method of estimation (e.g. non-linear least squares). For a detailed discussion, see Greene (1990), pgs.239-276.

$$P_i = f(X_i, Z_i)$$

where *P* is the price of the computer system *i*, *X* represents the set of explanatory variables which may or may not be linear and *Z* is the set of explanatory variables which by definition are linear (dummy variables and variables with non-positive values). In this paper, several functions were considered, these being linear and log-linear models (semi and double-log), models using a Box-Cox transformation and the non-linear model proposed by Triplett (1987).

3.1 - Linear and Log-Linear Models

The linear, semi-log or the double-log functional forms evaluated in this study are presented below.

Model 1. Linear - $P_i = \beta_0 + \sum_{k=1}^{n} \beta_k X_{ik} + \sum_{j=1}^{l} \gamma_j Z_{ij} + e_i$, where $e_i \sim NID(0,\sigma^2)$ (1)

Model 2. Semi-log - $\ln P_i = \beta_0 + \sum_{k=1}^{n} \beta_k X_{ik} + \sum_{j=1}^{l} \gamma_j Z_{ij} + e_i$, where $e_i \sim NID(0,\sigma^2)$ (2)

Model 3. Double-log - $\ln P_i = \beta_0 + \sum_{k=1}^{n} \beta_k \ln X_{ik} + \sum_{j=1}^{l} \gamma_j Z_{ij} + e_i$, where $e_i \sim NID(0,\sigma^2)$ (3)

In the double-log version, the number of versions that can be estimated will depend on *k*.[10]

3.2 - Box-Cox Model

The choice between linear and log-linear functional forms is rather limited and other types of transformations are available. One general type transformation is the Box-Cox transformation, which has a greater degree of flexibility.[11] The Box-Cox transformation takes the form of

$$y^{(\lambda)} = \frac{y^{\lambda} - 1}{\lambda} \quad .$$

where *y* is the variable being considered. This is a very useful model because it covers a wide range of transformations. Firstly, values of 1 and 0 for λ result in linear and log values of the variable y (models (2) and (3)). However, when the value λ falls between 0 and 1, we have a non-linear result.[12] If we apply the Box-Cox transformation to (1), we get,

Model 4. Box-Cox - $P_i^{(\lambda)} = \beta_0 + \sum_{k=1}^{n} \beta_k X_{ik}^{(\lambda)} + \sum_{j=1}^{l} \gamma_j Z_{ij} + e_i$, where $e_i \sim NID(0,\sigma^2)$ (4)

which is the conventional Box-Cox model.[13] The Box-Cox model has several uses, one being to help arrive at or discriminate between the various functional forms (Darnell (1994)) and for this reason, the transformation was included in this analysis.

---

[10] In this analysis, estimates for two versions of model (3) are presented the general model (3A) and the model where RAM is the only explanatory variable transformed (3B).

[11] See Box and Cox (1964).

[12] The conventional Box-Cox model is not without its disadvantages, namely that the possible values of dependent variable are restricted. In addition, λ affects the properties of the residuals as well as the functional form of the equation (Davidson and Mackinnon (1993), pages 507 to 510). For these reasons, the results of the Box-Cox estimations were used only as a guide in determining whether sufficient evidence existed to support a non-linear model.

[13] See Davidson and MacKinnon (1993), pages 480 to 488.

3.3 - Non-linear Model

A second category of functions is the non-linear class of hedonic functions, of which the "t-identification" function, suggested by Triplett (1987) and tested in this study, is a special case. The "t-identification" model examined in this study is

Model 5. Non-linear - $\ln p_i = \beta_0 + \sum_{k=1}^{n} \beta_k X_{ik}^{\alpha_k} + \sum_{j=1}^{l} \gamma_j Z_{ij} + e_i$, where e_i~NID(0,σ^2) (5)

where β_k and α_k are > 0 for all k and again, Z is a set of linear variables.[14]

From a theoretical standpoint, this functional form is very appealing since it allows for hedonic contours, which bow out from the origin, whereas with the linear and semi-log functions the hedonic contours are linear and in the case of the double-log function, the contours actually bow in.

4. Results

4.1 Desktop Computers

In the case of desktop computers, the question of choosing a functional form applied to only four variables, these being PRICE, RAM, HD and SCORE, since they were the only non-dichotomous variables that were always positive.[15] The rest appear in the equations as linear variables. Several functional forms were evaluated and the criteria for comparing them were, the signs of the coefficients, comparison of appropriate "goodness of fit" statistics, and the values of coefficients.[16]

For the linear and log-linear models, the goodness of fit statistics considered were the log-likelihood functions (adjusted in the semi-log and double-log cases) and the Double-Length Artificial Regression Statistic (DLR) suggested by Davidson and MacKinnon (1993) as a test for functional form.[17]Out of the three functional forms tested (linear, semi-log and double-log), the double-log functional form with the log of the independent variable RAM fits the best (see Tables 4 to 7). In the linear case, a negative value was obtained for CD, while the log-likelihood statistic was the smallest of the three models. Furthermore, the DLR statistic strongly rejects the null hypothesis of a linear model against the general Box-Cox case. In the semi-log case, the log-likelihood statistic has improved markedly and the DLR statistic fails to reject the null of the transformation on the independent variable. Similarly, there is a failure to reject the null that λ=0 for the double-log model where only the log of RAM is considered. However, when the log of RAM, HD and SCORE were used, the DLR statistic rejects the null. From these results, it appears that all but one of the independent variables, RAM, should appear in the equation as linear variables, while the results also suggest that the dependent variable PRICE appear in its logarithmic form.

---

[14] This model differs slightly from Triplett's in that linear variables are included.

[15] The variable XCACHE contained several 0 values for desktop computers.
[16] Also, the residuals from all of the estimations for desktop and notebook computers were evaluated and tested for the presence of multicollinearity and heteroscedasticity with the conclusion that neither of these two conditions was present to any significant degree.
[17] The DLR statistic is essentially a χ^2 statistic with one degree of freedom and is used to choose between the linear and the semi and double–log models by testing them against the general Box-Cox model. There may be some instances where both models may seem reasonable and in addition to comparing values of their loglikelihood functions, this statistic can be used to choose among the models. In the linear case, the null hypothesis that λ = 1 is evaluated, whereas for the semi and double-log models, the null is that λ = 0. See Davidson and MacKinnon (1993), pages 502 to 504.

Since the linear and log-linear models are but two special cases of the Box-Cox transformation, other forms of the general Box-Cox model were estimated using non-linear least squares method to obtain the value of λ in (4), and the results are presented in Table 8. Several models were tested using the Box-Cox transformation, firstly where λ was restricted to 0 for P_i and then the unrestricted case, where a value for λ was estimated for both P_i and X_i. Of the restricted models (i.e. where $\lambda = 0$ for P) the model where only the variable RAM was subject to the Box-Cox transformation fits the best. The estimated value of λ is 0.34, indicating that the relationship between the log of P and RAM could be non-linear.[18]

When considering the non-linear 't-identification' function suggested by Triplett (1987), several versions of (5) were tested in order to determine which model fit best (see Tables 9 to 11). Based on the signs of coefficients, F-tests, Likelihood ratio tests and convergence results, the version with only RAM appearing as a non-linear variable had the best fit.[19] The estimated value for α is 0.34186, which satisfies the condition that α (and all other coefficients for that matter) > 0.

4.1 - Comparison of Estimated Models

The semi-log (2), double log (3B) and the non-linear model (5B) appear to be reasonable models although on the basis of fit, the non-linear model appears to have a slight advantage.[20] However, one very important criterion for discriminating among these functional forms is how each model performs when an actual quality change is assessed – i.e. the values of the estimated coefficients. If the coefficients for a particular model result in hedonic price changes that are inconsistent or out of line with what is observed in the industry, then this model should be discarded in favour of a more realistic one. For example, if doubling the amount of RAM results in a doubling in the price of a computer, then such a model should be rejected in favour of a model which produces more realistic changes in price.[21]

A comparison of prices obtained using the three estimated models is presented in Table 12. This comparison was based on a typical model change surrounding three common characteristics: an upgrade in RAM from 64 MB to 128 MB, an increase in the hard disk size (HD) from 8.4 GB to 12.9 GB, and a change in CPU chip (or processor class - SCORE) from Pentium II 300 to Pentium II 400. The hypothetical price of the computer before the upgrades was $2,000 Cdn.[22]

From the results, it is clear that the non-linear model does not provide reasonable estimates when compared to the other two models. Using model (5B), a new price of $6,060.18 was obtained, while prices $2,945.52 and $2,905.16 were obtained for (2) and (3B). The latter two prices are clearly more realistic and in choosing between these two models, (3B) is the better model based on all of the criteria mentioned.

---

[18] The case for rejecting the linear model is further supported by the fact that when λ is estimated for both P and RAM, (i.e. the unrestricted model), the results are more robust. The loglikelihood function has improved noticeably, though caution must be used when interpreting these results, since in the conventional Box-Cox model, λ affects the residuals as well as the functional form of the equation (see Davidson and Mackinnon, (1993) pages 507 to 510.). In addition, with the Box-Cox model, theoretically, one can estimate different values for λ for *each* variable, though this can cause problems in the actual estimation procedure (see White (1993) page 152). When this procedure was tried, the model failed to converge.

[19] Models with the variable HD appearing in non-linear form did not converge for the most part, except when the combination of RAM and HD was considered. After further verification, this was considered a local and not a global maximum, resulting in a withdrawal of this model (see White, Judge et al, (1988) page 505).

[20] The loglikelihood functions for (2) and (3B), estimated using non-linear least squares to be comparable to the non-linear model (5B), were 115. 0178 and 122.8366 respectively.

[21] According to current industry trends, the price of RAM in Canadian dollars is in the neighborhood of $3 per MB, excluding installation charges and taxes.

[22] The price with upgrades was calculated by multiplying the changes in each upgrade by the appropriate coefficient(s), summing the values and taking the antilogarithm to arrive at a price change factor. The original price was then multiplied by this factor to arrive at the new price.

4.3 Notebook Computers

Similar results were obtained for notebook computers, namely that the double log model, where only the log of RAM is included, fits the best (see Tables 13 -16). Based on the values of the estimated log-likelihood and DLR statistics, the linear model can be clearly rejected, leaving the semi-log (2) and both double-log versions (3A and 3B) for consideration.[23] Upon further examination of their respective log-likelihood functions, the model containing the log of RAM (3B) has the better fit.[24]

The results obtained for the Box-Cox estimations (see Table 17) also favour the double-log model (3B). Unlike in the case of desktop computers, there is very little evidence to support a non-linear model (i.e. $0 < \lambda < 1$). For the restricted model with $\lambda = 0$ for P, the estimated value of λ for RAM is also 0. Therefore, it is no coincidence that poor results were obtained for the 't-identification' model, namely that a unique value for all of the coefficients could not be calculated.[25]

4.4 - Comparison of Estimated Models - Notebooks

As in the case of desktops, one additional consideration for evaluating these models is to compare their performance when the characteristics of a notebook computer change. The results from conducting such a hypothetical model change are presented in Table 18. This comparison was based on: an upgrade in RAM from 32 MB to 64 MB, an increase in the hard disk size (HD) from 3.2 GB to 4.0 GB, and a change in CPU chip (or processor class - SCORE) from a Pentium 233 MMX to a Pentium II 266. The hypothetical price of the computer before the upgrades was $2,900 Cdn.[26]

Not surprisingly, the resulting price changes for the double-log versions (3A and 3B) are similar (a difference of approximately $98). The semi-log (2) price change is noticeably higher when compared to that of (3A) - a difference of approximately $219, and similar to that of (3B) - a difference of roughly $121. Given the fact that model with the log of RAM (3B) has the better fit, this model is advocated over the semi-log version (2).

5. Conclusion

[23] For the linear, semi-log and double-log models all estimated coefficients have the correct sign, including WGHT, which should be inversely related to PRICE.

[24] In fact, by comparing the results for models 3A and 3B, it is apparent that using the log form of the variables HD, SCORE, XCACHE, WGHT and DSIZE does not improve the fit of the model at all. On the basis of parsimony, the simpler model (3B) is preferred.

[25] Specific results for the non-linear case were not worth including in this paper, sufficed to say that numerous models failed to converge, and for those that did converge, a local maximum was found producing nonsensical coefficient values.

[26] The price with upgrades was calculated in the same manner as for desktop computers.

In this study a variety of functional forms, ranging for linear to non-linear, were estimated, analysed and compared in order to arrive at a form which fits best and provides realistic results. For desktop and notebook computers, the double-log model (where RAM is the only explanatory variable transformed) appears to fulfil these criteria the best. In closing, approximately three years have passed since the last revision of the hedonic regression equations for computers was carried out at Statistics Canada. Given the dynamic nature of this industry, clearly one further consideration for improving the process of quality adjustment, not discussed in this analysis but nevertheless underlying all of this work, is the frequency of updating these equations. In the future, it is the intention to have these hedonic regressions revised on a more frequent basis in an attempt to capture the vast technological changes occurring in this industry.[27]

[27] At the minimum, annual revisions are planned.

Table 1: Synopsis of hedonic equations estimated by Prices Division

	First Estimation		**Second Estimation**				**Third Estimation**	
Date	1989/90		1991				1995/96	
Data Source	Data Pro Reports		Data Pro Reports				Public Works and Government Services Micro Acquisition Guide	
Computer type	16 bit and 32 bit		16 bit and 32 bit				Pentium desktop	
Number of observations	185		89 and 64 or 153 total				41	
Functional Form	Semi-log		Linear				Semi-log	
Estimated model for Minicomputer	*Variable*	*Estimated Coefficient*	*Variable*	16 bit *Estimated Coefficient*	32 bit *Estimated Coefficient*	Combined *Estimated Coefficient*	*Variable*	*Estimated Coefficient*
	Constant	5.968					Constant	7.318
	Execution speed	0.037	Execution speed	100.1	246.0	179.3	Hard drive	0.000058345
	RAM	0.069	RAM	0.914	0.837	0.88	RAM	0.013819
	Hard Disk Size	0.032	Hard Disk Size	0.026	0.017	0.02	Speed	0.003407
	Internal drive bays	0.036	Avg. Access Time	n/a	-12.48	n/a	17” monitor	0.212789
	Company	0.277	No. of Disk Drives	n/a	n/a	594.7	Company	0.212789
	Mail order	-0.563	Serial Ports	-954.8	n/a	-694.3		
	Available expansion slots	0.024	Available expansion slots	215.6	385.3	281.3		
	Bit size	0.211	Task capability	n/a	n/a	-1614.2		
			Service/reliability	408.7	n/a	n/a		
	R^2	Over 0.80	Adjusted R^2	0.86	0.77	0.91	Adjusted R^2	0.78

* n/a – not applicable.

Table 2: List of variables - desktop computers

Variable	Source	Value	Variable Name
A) Quantitative variables			
1. RAM	IDC	Numeric value (MB) e.g., 16, 32,	RAM
2. CPU score	CPU Scorecard	Numeric value (score) e.g. 440, 494, 588, (see appendix A.)	SCORE
3. Hard drive	IDC	Numeric value (MB) e.g., 2000, 2100,	HD
4. Extended Cache	"	Numeric value (KB) e.g., 0, 256, 512,	XCACHE
B) Dummy Variables			
5. Hard drive type	"	If HDTYPE = UWSCSI3, then HDTYPU=1, 0 otherwise	HDTYPU
		If HDTYPE = IDEenh, then HDTYPID=1, 0 otherwise	HDTYPID
		If HDTYPE = other, then HDTYPU=0 and HDTYPID=0.	
6. CD	"	No = 0 Yes = 1	CD
7.Network card	"	No = 0 Ethernet = 1	NTWCD
8. SCSI Control	"	No = 0 Yes = 1	SCSI
9. Modem	"	No = 0 Yes = 1	MDM
10. Case type	"	Desk and Mid-tower = 0 Workstation = 1	CASE
11. Vendor	"	If VENDOR = Hewlett-Packard or Compaq ,then BQUAL=1 (BQUAL = Best Quality); ELSE BQUAL=0;	BQUAL

Table 3: List of variables - notebook computers

Variable	Source	Value	Variable Name
A) Quantitative variables			
1. RAM	IDC	Numeric value (MB) e.g., 16, 32.	RAM
2. CPU score	CPU Scorecard	Numeric value (scope) e.g. 440, 494, 588. (see appendix A.)	SCORE
3. Hard drive	IDC	Numeric value (MB) e.g., 2000, 2100.	HD
4. Extended Cache	"	Numeric value (KB) e.g., 0, 256, 512.	XCACHE
5. Weight	"	Numeric value (pounds) e.g., 2.0, 4.3, 7.9.	WGHT
6. Display size	"	Numeric value (inches) e.g. 12.1, 13.3.	DSIZE
B) Dummy Variables			
7. Monitor quality	"	DS = 0 (Dual-scan, low quality) AM = 1 (Active matrix, high quality)	DQUAL
8. Vendor	"	If VENDOR = IBM, then HIEND=1 (HIEND = High-End models); ELSE HIEND=0;	HIEND

Table 4: Desktops - Results for linear model (1)

Variable	Estimate	Probability Value
Intercept	-350.10	0.099
RAM	20.4200	0.000
HARD	0.04514	0.041
SCORE	0.33720	0.114
XCACHE	1.03360	0.000
CASE	393.820	0.005
CD	-13.266	0.898
MDM	67.4540	0.751
NTWCD	50.3220	0.576
SCSI	977.230	0.000
BQUAL	432.430	0.000
HDTYPU	1502.20	0.000
HDTYPID	588.200	0.000
$R^2 =$ 0.7787		R^2 adj. = 0.7740
$\sigma =$ 946.33		
Log-likelihood = -4757.83		
Davidson/MacKinnon DLR statistic - $\chi^2_{(1)} =$		495.8390

Table 5: Desktops - Results for semi-log model (2)

Variable	Estimate	Probability Value
Intercept	6.678100000	0.000
RAM	0.003390100	0.000
HARD	0.000023039	0.000
SCORE	0.000250930	0.000
XCACHE	0.000585340	0.000
CASE	0.171610000	0.000
CD	0.053291000	0.015
MDM	0.055399000	0.219
NTWCD	0.032204000	0.092
SCSI	0.346070000	0.000
BQUAL	0.105000000	0.000
HDTYPU	0.312810000	0.000
HDTYPID	0.171750000	0.000
$R^2 =$ 0.8550		R^2 adj. = 0.8519
$\sigma =$ 0.20045		
Log-likelihood = -4416.09		
Davidson/MacKinnon DLR statistic - $\chi^2_{(1)} =$		0.06114

Table 6: Desktops - Results for double-log model (3A)

Variable	Estimate	Probability Value
Intercept	3.974000000	0.000
LRAM	0.355170000	0.000
LHARD	0.119800000	0.000
LSCORE	0.139720000	0.005
XCACHE	0.000585557	0.000
CASE	0.165080000	0.000
CD	0.018118000	0.417
MDM	0.077837000	0.085
NTWCD	0.024055000	0.209
SCSI	0.247370000	0.000
BQUAL	0.099690000	0.000
HDTYPU	0.356850000	0.000
HDTYPID	0.195650000	0.000
R^2 =	0.8551	R^2 adj. = 0.8520
σ =	0.20039	
Log-likelihood =	-4415.92	
Davidson/MacKinnon DLR statistic - $\chi^2_{(1)}$ =		10.2105

Table 7: Desktops – Results for double-log model (3B) RAM only.

Variable	Estimate	Probability Value
Intercept	5.641200000	0.000
LRAM	0.342100000	0.000
HARD	0.000019173	0.000
SCORE	0.000188440	0.000
XCACHE	0.000569590	0.000
CASE	0.161490000	0.000
CD	0.019959000	0.360
MDM	0.070749000	0.112
NTWCD	0.022961000	0.224
SCSI	0.276120000	0.000
BQUAL	0.106460000	0.000
HDTYPU	0.348220000	0.000
HDTYPID	0.201540000	0.000
R^2 =	0.8589	R^2 adj. = 0.8559
σ =	0.19774	
Log-likelihood =	-4408.27	
Davidson/MacKinnon DLR statistic - $\chi^2_{(1)}$ =		0.8053

Table 8: Desktops - Results for Box –Cox estimations of model (4)

Model		Estimated λ	Log-likelihood for Estimated λ	Log-likelihood for λ =0
Unrestricted	P and RAM	-0.36	-4379.92	-4408.27
Restricted (λ = 0 for P)	RAM, HARD and SCORE	0.47	-4405.84	-4415.92
	RAM and HARD	0.40	-4404.04	-4410.85
	RAM	0.34	-4403.88	-4408.27
	HARD	-0.06	-4413.93	-4413.94
	RAM and log of HARD	0.41	-4404.60	-4410.85

Table 9: Desktops – Results of various non-linear model estimations.

1.) Convergence of models	
Model	*Results*
RAM + HARD + SCORE	No convergence after 1500 iterations
RAM + HARD	Convergence after 101 iterations
RAM	Convergence after 148 iterations
HARD	No convergence after 1500 iterations
2. F-test against non-linear model of RAM	
Model	*Statistic*
RAM + HARD+ SCORE	-11.6212
RAM + HARD	-0.1811
3. Likelihood ratio test against non-linear model of RAM	
Model	*Statistic*
RAM + HARD+ SCORE	-5.9324
RAM + HARD	-.01752

Table10: Desktops– Results for non-linear model (5A)

Variable	**Estimate**	**Probability Value**
Intercept	262.68000	0.418
RAM - β_{Ram}	0.17436000	0.159
α_{Ram}	0.38563000	0.000
HARD - β_{Hard}	0.00002217	0.681
α_{Hard}	0.99131000	0.000
SCORE - β_{SCORE}	-257.540000	0.427
α_{SCORE}	-0.68669000	0.431
XCACHE	0.00058533	0.000
CASE	0.155610000	0.000
CD	0.030314000	0.164
MDM	0.067347000	0.133
NTWCD	0.026200000	0.053
SCSI	0.284390000	0.166
BQUAL	0.107780000	0.000
HDTYPU	0.329090000	0.000
HDTYPID	0.189860000	0.000
σ =	0.19502	
Loglikelihood =	124.2577	
SSE =	21.906	

Table 11: Desktops – Results for non-linear model (5B)

Variable	Estimate	Probability Value
Intercept	6.05590000	0.000
RAM - β	0.23419000	0.200
α	0.341860000	0.001
HARD	0.000018943	0.000
SCORE	0.000199520	0.000
XCACHE	0.000577370	0.000
CASE	0.153620000	0.000
CD	0.031326000	0.152
MDM	0.065922000	0.136
NTWCD	0.025065000	0.163
SCSI	0.300520000	0.000
BQUAL	0.108840000	0.000
HDTYPU	0.322470000	0.000
HDTYPID	0.191910000	0.000

σ = 0.19402
Log-likelihood = 127.2194
SSE = 21.682

Table 12: Desktops -Comparison of price changes using estimated models.

Model	Starting Price ($)	Price With Upgrades ($)	% change
1. Semi-log (2)	2,000	2,945.52	47.3
2. Double-log with log of RAM (3B)	2,000	2,905.16	45.3
3. T-identification (5B)	2,000	6,060.18	203.0

Upgrades;

1. *ΔRAM*. Up to 128 MB from 64 MB = 64 MB.
2. *ΔHARD*. Up to 12.9 GB from 8.4 GB = 4.5 GB.
3. *ΔSCORE*. Up to 1130 (Pentium II 400) from 865 (Pentium II 300) = 265.

Table13: Notebooks – Results for linear model (1).

Variable	Estimate	Probability Value
Intercept	-3485.8	0.000
RAM	6.45120	0.006
HARD	0.19805	0.000
SCORE	1.29620	0.001
XCACHE	2.15810	0.000
WGHT	-179.33	0.000
DSIZE	365.490	0.000
DQUAL	228.900	0.172
HIEND	653.430	0.000
$R^2 =$ 0.8058		R^2 adj. = 0.7964
$\sigma =$ 525.21		
Loglikelihood = -1332.17		
Davidson/MacKinnon DLR statistic - $\chi^2_{(1)} =$		6.76084

Table 14: Notebooks – Results for semi-log model (2).

Variable	Estimate	Probability Value
Intercept	6.176600000	0.000
RAM	0.001554800	0.000
HARD	0.000039409	0.000
SCORE	0.000427790	0.000
XCACHE	0.000602900	0.000
WGHT	-0.054339000	0.000
DSIZE	0.104110000	0.009
DQUAL	0.105920000	0.000
HIEND	0.156210000	0.000
$R^2 =$ 0.8215		R^2 adj. = 0.8129
$\sigma =$ 0.12655		
Loglikelihood = -1319.29		
Davidson/MacKinnon DLR statistic - $\chi^2_{(1)} =$		0.04535

Table 15: Notebooks – Results for double-log model (3A).

Variable	Estimate	Probability Value
Intercept	-0.05425000	0.911
LRAM	0.117630000	0.001
LHARD	0.257100000	0.000
LSCORE	0.291580000	0.000
LXCACHE	0.176980000	0.000
LWGHT	-0.232140000	0.000
LDSIZE	1.140200000	0.000
DQUAL	0.093781000	0.024
HIEND	0.156650000	0.000
$R^2 =$	0.8157	R^2 adj. = 0.8067
$\sigma =$	0.12861	
Loglikelihood =	-1322.10	
Davidson/MacKinnon DLR statistic - $\chi^2_{(1)} =$		0.04461

Table 16: Notebooks – Results for double-log model (3B) RAM only.

Variable	Estimate	Probability Value
Intercept	5.893600000	0.000
LRAM	0.113760000	0.006
HARD	0.000038954	0.000
SCORE	0.000422990	0.000
XCACHE	0.000605140	0.000
WGHT	-0.05049700	0.000
DSIZE	0.096929000	0.000
DQUAL	0.103450000	0.000
HIEND	0.158190000	0.000
$R^2 =$	0.8245	R^2 adj. = 0.8160
$\sigma =$	0.12548	
Loglikelihood =	-1317.82	
Davidson/MacKinnon DLR statistic - $\chi^2_{(1)} =$		0.02786

Table 17: Notebooks – Results for Box –Cox estimations of model (4).

Model		Estimated λ	Log-likelihood for Estimated λ	Log-likelihood for λ =0
Unrestricted	P only (semi-log)	-0.04	-1319.26	-1319.29
	P and RAM, HARD, XCACHE, SCORE, WGHT, DSIZE (double-log model 3A)	0.04	-1322.08	-1322.10
Restricted (λ = 0 for P)	RAM (double log model 3B)	0.00	-1317.82	-1317.82

Table 18: Notebooks – Comparison of price changes using estimated models.

Model	Starting Price ($)	Price With Upgrades ($)	% change
4. Semi-log (2)	2,900	3,715.21	28.1
5. Double -log (3A)	2,900	3,496.13	20.6
6. Double-log with log of RAM (3B)	2,900	3,593.78	23.9

Upgrades;

4. *ΔRAM*. Up to 64 MB from 32 MB = 32 MB.
5. *ΔHARD*. Up to 4.0 GB from 3.2 GB = 0.8 GB.
6. *ΔSCORE*. Up to 793 (Pentium II 266) from 546 (Pentium 233 MMX) = 247.

Appendix A -Scores of CPU processors

CPU	SCORE
1. Pentium 150 MMX	398
2. Pentium 166 MMX	440
3. Pentium 200 MMX	494
4. Pentium Pro 200	588
5. Pentium 233 MMX	546
6. Pentium 266 MMX	634
7. Pentium 300 MMX	709
8. Pentium Celeron 266	571
9. Pentium Celeron 300	603
10. Pentium Celeron 300A	776
11. Pentium Celeron 333	830
12. Pentium Celeron 366	998
13. Pentium Celeron 400	1092
14. Pentium Celeron 433	1183
15. Pentium II 266PE	815
16. Pentium II 300PE	879
17. Pentium II 233	704
18. Pentium II 266	793
19. Pentium II 300	865
20. Pentium II 333	949
21. Pentium II 350	1000
22. Pentium II 366	1053
23. Pentium II 400	1130
24. Pentium II 450	1240
25. Pentium II Xeon 450	1370*
26. Pentium III 450	1500
27. Pentium III 500	1650
28. Pentium III Xeon 500	1815*
29. Pentium III Xeon 550	1997*

*For the Xeon series, the scores were estimated using the benchmark data available on the Internet site 'www.intel.com' and communication from the people responsible for scoring the various CPUs on 'The CPU Scorecard' Internet site.

References

Berndt, Ernst R. (1991), *The Practice of Econometrics: Classic and Contemporary,* Addison-Wesley Publishing Company, Reading Massachusetts.

Box, G.E.P. and D. Cox (1964), 'An Analysis of Transformations', *Journal of the Royal Statistical Society,* Series B, 211-264.

Darnell, Adrian C. (1994), *A Dictionary of Econometrics,* Vermont: Edward Elgar Publishing Limited.

Davidson, R., and J.G. MacKinnon (1993), *Estimation and Inference in Econometrics*, New York: Oxford University Press.

Greene, William H. (1990), *Econometric Analysis,* New York: Macmillan Publishing Company.

Judge, G. G., R.C. Hill, W.E. Griffiths, H. Lutkepohl and T.Lee (1988), *Introduction to the Theory and Practice of Econometrics,* New York: John Wiley & Sons Inc.

Microsoft Press (1991) *Computer Dictionary*, Macmillan of Canada.

PC WORLD, (1998), "Problem-free PCs - Who Makes Them?", San Francisco, November.

PC WORLD, (1996), "Best PCs for Windows 95", San Francisco, March.

Triplett, Jack E. (1987), "Price and Technological Change In A Capital Good: A Survey Of Research On Computers" *Bureau of Economic Analysis, Discussion Paper 23*.