NBER WORKING PAPER SERIES

PRICE INDEXES FOR MICROCOMPUTERS: AN EXPLORATORY STUDY

Ernst R. Berndt

Zvi Griliches

Working Paper No. 3378

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 1990

We gratefully acknowledge the financial support of the Alfred P. Sloan Foundation to the NBER. We have benefited from the able research assistance of Joanna Stavins, and by having access to earlier research on this topic by Jeremy Cohen and Amy Kim. Proprietary data on micro-computer shipments and installations by PC model were kindly provided us by Bruce A. Stephen and Lloyd Cohen of the International Data Corporation. This paper is part of NBER's research program in Productivity. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Working Paper #3378
June 1990

PRICE INDEXES FOR MICROCOMPUTERS: AN EXPLORATORY STUDY

ABSTRACT

In this paper we focus on alternative procedures for calculating and interpreting quality-adjusted price indexes for microcomputers, based on a variety of estimated hedonic price equations. Our data set comprises an unbalanced panel for 1265 model observations from 1982 to 1988, and includes both list and discount prices. We develop and implement empirically a specification test for selecting preferable hedonic price equations, and consider in detail the alternative interpretations of dummy variable coefficients having time and age, vintage and age, and all of the time, age, and vintage dummy variables as regressors.

We then calculate a variety of quality-adjusted price indexes; for the Divisia indexes we employ estimated hedonic price equations to predict prices of unobserved models (pre-entry and post-exit). Although our indexes show a modest amount of variation, we find that on average over the 1982-88 time period in the US, quality-adjusted real prices for microcomputers decline at about 28% per year.

Ernst R. Berndt
NBER
1050 Massachusetts Avenue
Cambridge, MA 02138

Zvi Griliches
NBER
1050 Massachusetts Avenue
Cambridge, MA 02138

I. INTRODUCTION

In recent years, a considerable amount of research has focussed on the construction and interpretation of price indexes for computers.[1] The computer market is a fascinating one, for technological change has been rapid, there has been a great deal of entry and exit of firms and models, and particularly in the microcomputer market, models have been simultaneously sold at different prices by standard retail and discount vendors.

Because of the rapid technological change and turnover of models and firms, Franklin M. Fisher et al. [1983] have characterized the mainframe computer market as typically being in "disequilibrium". One consequence of this is that price indexes have been used for two rather different purposes, one to deflate expenditures or purchases into constant dollars, and the other to trace out movements in a technological frontier, such as a price-performance ratio.

If quality-adjusted prices reacted instantaneously and fully to the introduction of new technology, then an index that traced out the technological frontier would be identical to one that covered all models sold in the marketplace. In periods of disequilibrium, however, the two quality-adjusted price indexes might differ, with consumers tolerating transactions at more than one quality-adjusted price. Reasons for such multiple price transactions include the fact that the supply of some new models might initially be limited, and that in spite of this excess demand, manufacturers may offer new models at lower prices to facilitate dissemination of information about the impact of the new low-price technology. Alternatively, surviving models may be of higher quality in some unobserved characteristics, or may benefit from the accumulation of specialized software and knowhow. The extent of such price disequilibrium is of course an empirical issue; some recent evidence on this issue for the mainframe computer market is presented by Ellen Dulberger [1989].

Although the mainframe computer market has received considerable attention, to the best of our knowledge there has been little empirical work on the microcomputer or PC market.[2] In this paper we focus attention on the interpretation of implicit price indexes and coefficients from hedonic price equations using detailed data from the retail and discount US microcomputer (PC) markets.[3] Our work builds on the research of two of our students, Jeremy Cohen [1988] and Amy Kim [1989]. Cohen originally gathered and assembled price and characteristics data covering the 1976-1987 time period; the data he updated to include 1988 were then examined further by Kim.

Based on hedonic regression equations with pooled data, both Cohen and Kim generated implicit PC price indexes for retail (list) and discount markets. Representative findings from Cohen and Kim are presented in Table 1 below, as are the PC price indexes computed by Robert J. Gordon [1990] based on 21 PC model-years and the "matched model" index number procedure, and the BEA "official" PC price index.[4]

As is seen in Table 1, all indexes suggest rapid declines in the quality-adjusted price of microcomputers. Cohen reports an average annual growth rate (AAGR) of -25.36% in the <u>real</u> (CPI adjusted) price of PC's over the 1976-87 time frame for list prices, and a slightly lower -21.33% for discount prices. Kim finds an AAGR of -29.48% for real list prices 1976-88, and -23.53% for real discount prices. Gordon's calculations suggest an AAGR

Table 1

ALTERNATIVE PRICE INDEXES FOR MICROCOMPUTERS

Year	Hedonic Regression Equations: Cohen [1988] List (Real)	Cohen [1988] Discount (Real)	Kim [1989] List (Real)	Kim [1989] Discount (Real)	Matched Model Procedure Gordon [1989] (Real)	Gordon [1989] (Nominal)	BEA "Official" PC Price Index (Real)	BEA Index (Nominal)	CPI
1976	4.7709								.5828
1977	2.7347								.6262
1978	2.0878	1.4558							.6727
1979	1.8015	1.3638							.7471
1980	1.6923	1.4726							.8535
1981	1.4189	1.2700			1.3441	1.2561			.9345
1982	1.0000	1.0000	1.000	1.000	1.0000	1.0000	1.000	1.000	1.0000
1983	.7118	.4613	.687	.464	.7459	.7687	.777	.801	1.0306
1984	.5926	.6225	.617	.920	.5576	.5939	.568	.605	1.0651
1985	.3898	.3798	.409	.595	.3871	.4287	.511	.566	1.1076
1986	.2581	.2494	.268	.393	.2916	.3292	.369	.417	1.1291
1987	.1913	.1680	.194	.259	.2201	.2579	.236	.276	1.1715
1988			.123	.200					1.2176
1982-87 AAGR:	-28.16%	-30.01%	-27.96%	-23.68%	-26.12%	-23.74%	-25.08%	-22.70%	3.22%

Notes: Data are taken from Cohen [1988], Appendix D, p. 70 (renormalized to 1982 = 1.000); from Kim [1989], Appendix 22; and from Gordon [1990], Table 6.13, p. 237 (renormalized). The BEA Index is from the November 1988 issue of the Survey of Current Business. Gordon's index is based on data covering 21 model-years observed over the 1981-87 time period in advertisements in Business Week and PC Magazine. The real price indexes are computed as the nominal price index divided by the Consumer Price Index (the CPI is in the last column).

of -26.12% for the shorter 1982-87 time period (a mix of list and discount prices), while the BEA real price index falls at an AAGR of -25.08% for the same time period.

To facilitate comparison of indexes, in the bottom row of Table 1 we present AAGR for all the real and nominal price indexes over the same 1982-87 time interval. Note that the BEA and Gordon price indexes based on matched model procedures show a less rapid decline over this interval than do the hedonic regression indexes constructed by Cohen and Kim.

The research results we report in this paper extend the work of Cohen, Kim and Gordon in a number of related ways. First, we focus attention on the more general interpretive implications of the fact that the PC market is a changing one during the 1982-88 time period, involving shake-outs of some models, successful innovations for others, and dramatic changes in product characteristics. The data sample we observe is opportunistic in the sense that it represents only new and surviving vintages. We examine whether surviving vintages are priced at a premium, and how prices of surviving vintages adjust when new models are introduced incorporating technological advances.

Second, we examine several econometric implications of the fact that our data are in the form of an unbalanced panel, due to differential survival rates in the marketplace. In particular, we explore implications for estimation of how one implements empirically the identity that the year in which the model is observed is the sum of the year in which the model was first introduced and its age in years. A diagnostic test is discussed and implemented for checking our hedonic specification. Issues of sample selectivity are also addressed.

Finally, we construct and comment on a variety of price indexes that attempt in various ways to serve as deflators, or as measures that trace out a technological frontier in the PC market.

II. THE DATA

The data set available for this study includes price and technical attribute measures for new (not used) personal computers sold in the US from 1976 to 1988. The 1976-1987 data were originally collected and analyzed by

Jeremy Cohen [1988]; these data have been updated by Cohen to 1988, they have been employed by Amy Kim [1989] in further analysis, and have undergone additional revisions by us. The primary source of technical data was the Byte magazine comprehensive technical reviews. Since both list and discount prices often varied within each calendar year, the June issues of Byte, PC Magazine, and PC World were employed for list price data, while ads in the Science and Technology, and Business Sections of a Sunday New York Times issue in early June of each year were employed to obtain discount prices.[5] Additional data sources included the Dataquest Personal Computer Guide, and IBM pricing and technical data.

Characteristic and performance data collected by Cohen include RAM (the amount of random access memory standard on each PC model, measured in kilobytes), MHZ (the clock speed, measured in megahertz, or millions of cycles per second), HRDDSK (the amount of storage on the hard disk, if one exists, measured in megabytes), NUMFLP (the number of floppy disk drives standard on each model), SLOTS (the total number of eight, sixteen and thirty-two bit slots available for expansion boards), and AGE (the number of years the model has been sold on the market, where the model has an age of zero in its initial year).

As we noted earlier, an important feature of the PC market is that it is changing very rapidly. A model introduced in year zero may survive with unchanged characteristics into year one, two, or even longer, or, as is often the case, it may survive with differing characteristics into other years (we call this a changed version of the model). Other models may exit after being in the market only one year. Hence the stock of models sold in any given year consists of new and incumbent models, and among the incumbent models, new and old versions.

To highlight the evolution of the PC market, in Table 2 we present arithmetic means of characteristics for models newly introduced from 1982 through 1988. As is seen there, the mean nominal price decreased slightly, about 3% from $3617.61 in 1982 to $3508.47 in 1988, while mean RAM increased more than tenfold from 94.92 to 1069.39 KB, MHZ clock speed jumped more than three times from 4.4046 to 14.8201, and the mean hard disk storage rose from 0 to 43.638 MB.

Table 2

MEAN VALUES OF CHARACTERISTICS FOR NEW MODELS, 1982-88

YEAR	N	RAM	MHZ	HRDDSK	NUMFLP	SLOTS	NOMINAL PRICE
1982	13	94.92	4.4046	0.000	1.154	3.308	3617.61
1983	59	122.78	4.6807	2.161	1.237	3.322	3017.66
1984	80	204.00	5.1998	3.012	1.338	3.325	3026.96
1985	61	326.69	5.9974	4.607	1.295	4.000	2991.15
1986	123	539.25	7.6016	11.220	1.195	5.081	2955.60
1987	245	773.09	10.1033	22.355	1.098	5.016	3251.40
1988	141	1069.39	14.8201	43.638	1.014	5.993	3508.47

Note: N is the number of new models by year. Other variables are defined in the text.

Although newly introduced models constitute a major portion of our PC data base (722 of the 1265 models are new), some models last several years, and of these incumbent models, some evolve into different versions with altered characteristics. Other models exit, leaving the marketplace after less than one year. In Table 3 we summarize the mixed nature of the PC market from 1982 to 1988, including new and up to age 3 models, separately for the total, retail and

Table 3

VINTAGE COMPOSITION OF MICROCOMPUTER MARKET
1982-88

NUMBER OF PRICE OBSERVATIONS

YEAR		AGE - 0	AGE - 1	AGE - 2	AGE - 3	TOTAL
1982	Total	13	7	12	9	41
	Retail	10	5	4	5	24
	Discount	3	2	8	4	17
1983	Total	59	9	5	6	79
	Retail	53	5	3	1	62
	Discount	6	4	2	5	17
1984	Total	80	44	3	0	127
	Retail	63	25	2	0	90
	Discount	17	19	1	0	37
1985	Total	61	39	12	2	114
	Retail	59	18	5	0	82
	Discount	2	21	7	2	32
1986	Total	123	35	23	6	187
	Retail	106	26	13	2	147
	Discount	17	9	10	4	40
1987	Total	245	92	42	11	390
	Retail	217	63	30	9	319
	Discount	28	29	12	2	71
1988	Total	141	146	32	8	327
	Retail	129	59	5	0	193
	Discount	12	87	27	8	134
Grand Total		722	372	129	42	1265
Retail Total		637	201	62	17	917
Discount Total		85	171	67	25	348

discount markets. For the total market, 58% (722 of 1265) are new models, 29% (372) models survived one year (perhaps with changed characteristics and reduced prices to meet the market competition from frontier models), 10% (129) survived two years, and 3% (42) remained in the market place for three years.

Altogether, about 72% of our model observations are taken from the retail (list price) market, while 28% represent discount (much of it mail order) quotations.[6] However, as is also seen in Table 3, the age composition of models varies considerably between the retail and discount markets. Specifically, discount markets tend to have a much smaller proportion of new models, and much larger proportions of age 1, 2 and 3 year-old models. Finally, it is worth noting that in our data set, some models are sold in both the retail and discount markets (e.g., IBM and Compaq) and are therefore "observed" twice, while others are only in the retail market (e.g., PC Limited); however, no model is observed only in the discount market.

To focus attention on issues involved in interpreting coefficients with unbalanced panels, in this paper we adopt in essence the regressors and functional form employed by Cohen and Kim, in which the logarithm of the real price (LRPRICE) is regressed on the logarithm of several characteristics including LRAM, LMHZ, LHRDDSK (log (HRDDSK + 1)), LNUMFLP (log(NUMFLP + 1)), LSLOTS (log(SLOTS + 1)), and a number of dummy variables.

The dummy variables for characteristics include PROC16 (=1 if model has a 16-bit processor chip, else zero), PROC32 (=1 if model has a 32-bit processor chip), DBW (=1 if system comes with a monochrome monitor), DCOLOR (=1 if system comes with a color monitor), DPORT (=1 if model is portable or convertible), DEXTRA (=1 if model has a significant piece of additional hardware included, else zero. Examples of such extra hardware include modems, printers or an extra monitor), and DDISC (=1 if system price is discounted by the vendor).

The dummy variables for manufacturers are DIBM (=1 if system is made by IBM), DAPPLE (Apple), DCOMMO (Commodore), DCMPQ (Compaq), DNEC (NEC), DRDIOSH (Radio Shack), DPCLIM (PC Limited), and DOTHER (made or sold by any other company than those noted above).

Finally, a number of time and vintage effect dummy variables are employed. For time effects, the dummy variables T82, T83,...,T88, take on the value of one if the PC model was sold in that year, and else equal zero. For vintage effects, the dummy variables V79, V80,...,V88 take on the value one if the model was originally introduced in that year, and else equal zero.

The above variables, as well as several other measures, were included as regressors in a number of specifications examined by Cohen and by Kim using data beginning in 1976. Since the PC market was very small from 1976 until the entry of IBM in late 1981 (only 156 models were introduced before 1982), in this paper we confine our attention to the 1982-1988 time period and the 1979-88 vintages, restricting our sample to PC models whose age is three years or less, and dividing the AGE variable into three dummy variables, AGE1, AGE2 and AGE3, with a new model having an implicit age of zero. The data used in our regression analysis is summarized in Table 4, where we present sample means as well as minimum and maximum values of the various variables.

III. <u>ECONOMETRIC ISSUES</u>

Our data set comes in the form of an unbalanced panel, in that the number of observations by age, and by vintage, varies by year. Let the vintage of model i (the year in which it was first introduced) be V, where V = V79,V80,...,V88, let the year (time period) in which the model is observed be T, where T = T82, T83,...,T88, and let the age of the model of vintage V observed in time period T, in years, be A, where A = A0,A1,A2 and A3. This yields the identity that for any model-observation,

$$\text{TIME} = \text{VINTAGE} + \text{AGE}, \quad \text{or} \quad T = V + A. \tag{1}$$

Table 4

Summary of Microcomputer Data, 1982-1988

Variable	Mean	Minimum Value	Maximum Value
PRICE	2846.96	40.00	13995.00
RAM	560.73	1.00	4096.00
MHZ	8.3474	1.00	25.00
HRDDSK	17.199	0.00	314.00
NUMFLP	1.1526	0.00	2.00
SLOTS	4.5644	0.00	21.00
AGE	.5976	0.00	3.00
PROC16	.5510	0.00	1.00
PROC32	.1344	0.00	1.00
DBW	.4213	0.00	1.00
DCOLOR	.0285	0.00	1.00
DPORT	.1747	0.00	1.00
DEXTRA	.0206	0.00	1.00
DDISC	.2751	0.00	1.00
DIBM	.0988	0.00	1.00
DAPPLE	.0427	0.00	1.00
DCMDRE	.0285	0.00	1.00
DCMPQ	.0648	0.00	1.00
DNEC	.0427	0.00	1.00
DRDIOSH	.0490	0.00	1.00
DPCLIM	.0166	0.00	1.00
DOTHER	.6569	0.00	1.00
T82	.0324	0.00	1.00
T83	.0635	0.00	1.00
T84	.1004	0.00	1.00
T85	.0901	0.00	1.00
T86	.1478	0.00	1.00
T87	.3083	0.00	1.00
T88	.2585	0.00	1.00

Total Number of Observations: 1265

If T, V and A were continuous variables indexed by year, one could not simultaneously introduce all three as regressors in a linear equation to be estimated by least squares, for exact collinearity would result; to avoid such collinearity, only two of the three could be included directly, and estimates for the third could be computed indirectly using (1). Alternatively, as has been discussed by Stephen Fienberg and William Mason [1985], one could specify a model with nonlinear transformations of all three variables, such as their squared values.[7]

However, when T, V and A are discrete dummy variables rather than continuous, and if the coefficients of these variables are to be held constant over time and/or vintage, then the simple adding-up conditions implied by (1) no longer hold. This raises a number of issues involving the interpretation of dummy variable coefficients, and the maximal parameterization possible that avoids exact collinearity.[8]

Suppose one specified an equation having, in addition to an overall constant and a number of model i-specific characteristic variables (called X's), the time and age dummy variables (but not the vintage dummies),

$$\ln P_{ivat} = \alpha + T'\alpha_t + A'\alpha_a + X'\beta + u_{ivat}, \tag{2}$$

where the T82 and A0 dummy variables were deleted. In this case one might interpret estimates of the α_t as changes in the quality-adjusted price index relative to 1982, holding age fixed. Similarly, estimates of the α_a can be interpreted as the effects of age (relative to a new model of age zero) on price, holding time fixed. Intuitively, the α_t parameters in this T-A specification represent the general movement in average PC prices, given the average rate at which selectivity occurs in the sample.

While cumulated evidence from the mainframe market suggests that the α_t should decline with time,[9] it is not clear what one should expect for signs of the estimated α_a, which represent the effects of quality differentials on models of different ages sold contemporaneously, holding time fixed. According to one line of thinking, new models of superior quality should command a premium price, and if market competition forced the valuations of characteristics of incumbent models to obsolesce at the same rate, prices of surviving vintages would decline appropriately, and since the time dummy captures the full price effect, one might expect estimates of α_1, α_2 and α_3 to be approximately zero. On the other hand, selectivity in the marketplace reveals survival of the fittest, and if there are unmeasured characteristics (e.g., compatible software or hardware, differential service policies and warranties), then the age coefficients may to some extent be providing an estimate of the unobserved positive quality differentials among the survivors as vintages progress. To unscramble the obsolescence and selectivity components of the estimated age coefficients, one would need to assume that all of the quality differences among vintages were already captured in the changing computer characteristics and their associated coefficients, assuring thereby that age coefficients reflected selectivity alone.

Alternatively, one might specify a regression equation using the vintage V and age A dummy variables, rather than the T and A as in (2),

$$\ln P_{ivat} = \alpha + V'\alpha_v + A'\alpha_a + X'\beta + u_{ivat}, \tag{3}$$

where, say, the V82 vintage dummy variable was omitted. In terms of interpretation, note that if the technical characteristic variables captured virtually all of the quality changes embodied in models, then the vintage coefficients would essentially be capturing the average decline in prices by vintage, which in turn is some average of the difference between A and the implicit T over ages. Similarly, given that the specification (3) conditions on vintages, one can interpret the α_a age coefficients as representing some average of implicit time coefficients, since by the identity (1), A = T - V for all vintages.

In spite of its apparent similarity with (2) given the identity (1), the specification in (3) is in fact quite different, except for the special case when models of only one age are considered (e.g., only AGE=0 models). There are several reasons for this. First, the number of dummy variable coefficients is greater in (3) than in (2), for in (2) there are six time (T83-T88) and three age (A1-A3) coefficients, while in (3) there are nine vintage (V79-V88, V82 omitted) and three age (A1-A3) parameters. Thus in general one should not expect least squares estimates of α, α_a and β to be the same in the two specifications.

Second, while the age coefficients condition on time in (2), in (3) the conditioning is on vintage. In particular, in (3) the α_a age coefficients are interpreted as the effect on price of age (relative to a new model), holding vintage V (not time T) fixed. Since the conditioning changes, least squares estimates should also be expected to differ in (2) and (3).

To understand this better, consider a V-A specification such as (3). An implicit time coefficient such as that for, say, T88 could be computed in four different ways:

$$\alpha_{t,88} = \alpha_{v,88} + \alpha_{a,0}, \quad \alpha_{t,88} = \alpha_{v,87} + \alpha_{a,1},$$
$$\alpha_{t,88} = \alpha_{v,86} + \alpha_{a,2}, \text{ and/or } \alpha_{t,88} = \alpha_{v,85} + \alpha_{a,3}. \tag{4}$$

Similarly, with the T-A specification as in (2), the implicit vintage coefficient for, say, V85, could be computed alternatively as

$$\alpha_{v,85} = \alpha_{t,85} - \alpha_{a,0}, \quad \alpha_{v,85} = \alpha_{t,86} - \alpha_{a,1},$$
$$\alpha_{v,85} = \alpha_{t,87} - \alpha_{a,2}, \text{ and/or } \alpha_{v,85} = \alpha_{t,88} - \alpha_{a,3}. \tag{5}$$

Least squares estimation of the V-A and T-A specifications implicitly weight and average over these four possibilities in different ways, and thus there is no reason to expect implicit and direct estimates of the α_v, α_a, and/or α_t coefficients to be numerically equivalent in the T-A and V-A models.

But if the V-A and T-A specifications yield varying estimates because of their distinct conditioning and use of differing information, how is one to choose among them? For purposes of computing quality-adjusted price indexes, the directly estimated time coefficients based on (2) have a clear interpretation, and for that reason specification (2) has formed the basis of almost all hedonic price index studies. But is it necessary to delete the V variables completely -- can one not employ a specification that efficiently uses information simultaneously from the T, A and V dummy variables, yet avoids exact collinearity?

This issue has been addressed by Robert E. Hall [1971], whose context involved use of a balanced panel data set for second-hand trucks. In our context, the maximal parameterization consistent with avoiding exact collinearity among the T, V and A dummy variables turns out to be one in which eight of the original ten vintage dummy variables are added to the T-A specification (2), i.e. two (not one) of the vintage dummies are deleted from the original set of ten (V79 to V88).[10,11] We can write such a specification as

$$\ln P_{ivat} = \alpha + T'\alpha_t + A'\alpha_a + V'\alpha_v + X'\beta + u_{ivat}, \tag{6}$$

where the vector of dummy variables in V consists of eight elements. As Hall noted, coefficients on the α_v should be interpreted as <u>differences</u> from the average rate of growth of technical progress embodied but unobserved in pairwise comparisons of vintages. For example, if one omitted the V82 and V83 dummy variables, the α_v coefficient estimates should then be interpreted as a contrast, e.g., the coefficient on the V88 dummy variable should then be interpreted as the <u>difference</u> between the average 1988 vintage effect and the mean of the average vintage effects for 1982 and 1983. We suggest that a necessary condition for a hedonic price equation to be satisfactory is that the portion of quality change not captured by the characteristics variables should be unrelated to vintages, i.e. in a desirable specification, the α_v should be approximately zero.

It follows that since the α_v coefficients represent contrasts in average rates of growth due to unobserved quality change, one can interpret a test that the $\alpha_v = 0$ as corresponding to a test that changes in characteristics among models and over time adequately capture quality changes between vintages, for average unobserved vintage effects are not systematically different in pairwise comparisons among vintages. Further, if it were found that the α_v are simultaneously different from zero, then one might interpret that result as suggesting model misspecification, reflecting either the impact of omitted characteristic variables, or invalid stability constraints on the characteristics parameters over time. Hence, as noted above, a desirable specification would yield non-rejection of the null hypothesis that the α_v simultaneously equal zero, in which case (2) would be empirically supported as a special case of (6).[12]

Hypotheses concerning parameter restrictions can of course be tested using the standard F-test methodology. As has been emphasized by, among others, Kenneth Arrow [1960] and Makoto Ohta and Zvi Griliches [1976], when samples are large and standard test procedures are employed, one is likely to reject most simplifying parameter restrictions on purely statistical grounds, even though they may still serve as adequate approximations for the purpose at hand. There are several ways one can deal with this problem.

First, to accommodate the larger sample size, we can compensate by choosing very tight significance levels for the standard F-tests. In this paper we do that by choosing .01 significance levels. Second, one can adopt a more agnostic and conservative criterion that the null hypothesis holds only approximately rather than exactly in the sample. In such a case, as Edward Leamer [1978] has shown, one can employ a Bayes procedure that in essence decreases the significance level as the sample size n increases. Thus a second criterion we use to assess the validity of our hypotheses is the Bayes factor asymptotic approximation developed by Leamer [1978, pp. 108-114], translated from the condition that it exceeds one into an F-value expression. We call this Leamer-Bayes critical value the Bayes criterion.[13]

Third and finally, since our hedonic regressions are semi-logarithmic, the root mean squared errors (RMSE) measure the unexplained variation in prices in, roughly, percentage units. A reasonable criterion is to use the difference in the RMSE of the constrained and unconstrained regressions as a relevant measure of the price-explanatory power of a particular model. As our third test criterion, we therefore will reject the null hypothesis when the RMSE under the alternative results in a reduction of more than 5% in the RMSE (the standard deviation of the unexplained variation in log prices). With an average RMSE of around 0.40, this RMSE criterion implies that we are looking for a movement of at least about 0.02, say, 0.40 to 0.38, before we will "give up" on the more parsimonious parameterization implied by the null hypothesis.

IV. INITIAL RESULTS

We begin with results from a T-A model in which the time and age dummy variables are included, but the vintage dummies are omitted, as in (2). Results from such a regression are presented in Table 5. The dependent variable is the logarithm of the real price (LRPRICE), and the variables are essentially those as in Cohen and Kim. Regression results are reported for a pooled sample, retail price observations, and discount price observations. Recall that in many cases, a particular model appears in both the retail and discount markets. Given the specification of dummy variables, the estimated intercept term corresponds to that for a model of age zero in 1982 having an 8-bit processor, no monitor or extras, is not portable, is not in the discount market, and is made by IBM.

A number of results are worth noting. The coefficient on LMHZ is positive and significant in all three regressions, but is largest in the discount market; parameters on the LSLOTS variable follow a similar pattern. By contrast, parameters on LRAM and LNUMFLP, though positive and significant, are larger in the retail than discount market.

Second, in all three regressions the coefficients on the age variables are positive, holding time fixed, suggesting that the age effects of selectivity are substantial. Interestingly, the effect of age is largest in the retail market, where the age premium is statistically significant and increases with age, implying that list prices of all surviving computers do not drop "fast enough". In the discount market, however, the age coefficients are statistically insignificant and follow no pattern. This suggests that some type of selectivity is occuring in the transition from retail to discount markets, and that conditional on having entered the discount market, there is little age selectivity remaining.

In terms of other dummy variable coefficients, estimates of PROC16 and PROC32 are positive, statistically significant and larger in the retail than discount market, and the positive DPORT parameter estimate is larger in the discount market. Although the general pattern of the time dummy variable coefficient estimates is similar in the retail and discount markets -- both revealing declines in quality-adjusted prices since 1982, in the discount market the pattern of estimates between D83, D84 and D85 is not monotonic, suggesting that the discount market is more volatile, or that our discount sample is too small in these years to generate reliable parameter estimates.

Notice also that the DDISC coefficient in the pooled regression is negative (-.2903) and significant, as expected. We tested the null hypothesis that, aside from a parallel shift due to being in the discount market, all coefficients are identical in the retail and discount markets. The F-test statistic corresponding to this null hypothesis is 2.77, while the .01 critical value is 1.73. Hence, on the basis of the F-test criterion, the null hypothesis of parameter equality in discount and retail markets is rejected. However, the Bayes critical value is 7.39, considerably greater than the F-statistic of 2.77. In terms of RMSE, the improvement under the alternative hypothesis is 1.94%. Overall, we interpret these results as suggesting that the evidence supporting the null hypothesis of parameter equality (aside from a parallel shift) in the two markets is modest.

We also checked out two other Chow-type tests. First, we ran separate regressions for the age equal zero, one, two and three years old sub-samples, and compared the residual sums of squares with that from the pooled model reported in column one of Table 5. The calculated F-test statistic is 2.66, while the .01

Table 5

Results from Regression Models with Time and Age Dummy Variables Included
Data from 1982-88, Pooled, Retail and Discount Samples
Dependent Variable is LRPRICE

	Pooled Regression		Retail Models Only		Discount Models Only	
Variable	Estimated Parameter	t-Statistic	Estimated Parameter	t-Statistic	Estimated Parameter	t-Statstic
Intercept	4.8101	41.934	4.7316	33.152	4.4924	23.823
LRAM	.3140	14.804	.3313	12.746	.2721	7.875
LMHZ	.3157	7.668	.2197	4.409	.5482	7.620
LHRDDSK	.1688	19.876	.1716	17.710	.1543	9.005
LNUMFLP	.4304	8.588	.4753	7.869	.2913	3.365
LSLOTS	.1721	8.483	.1502	5.921	.2396	7.211
AGE1	.1193	3.911	.1296	3.531	.0414	0.735
AGE2	.1542	3.448	.2352	3.984	.0192	0.268
AGE3	.2984	4.034	.5333	4.748	.1469	1.454
PROC16	.2087	5.817	.2501	5.894	.1319	2.037
PROC32	.5193	8.101	.6560	8.829	.1926	1.500
DBW	.0261	0.844	.0222	0.633	-.0511	-0.944
DCOLOR	.0315	0.423	.0463	0.491	-.0129	-0.110
DPORT	.3565	8.943	.3400	6.763	.4703	7.273
DEXTRA	.2756	3.242	.2698	2.733	.4609	2.706
DDISC	-.2903	-9.460				
DAPPLE	.2729	3.627	.1982	1.999	.4470	3.938
DCMDRE	-.3291	-3.776	-.3763	-3.089	-.1226	-0.981
DCMPQ	.2678	4.176	.3598	4.045	.2266	2.394
DNEC	.1114	1.548	.2369	2.399	-.0265	-0.251
DRDIOSH	.0618	0.891	.0162	0.205	.4644	3.127
DPCLIM	-.5047	-4.927	-.4707	-4.402		
DOTHER	.0062	0.141	.0430	0.823	.0027	0.034
T83	-.3974	-4.768	-.2193	-2.081	-.8034	-5.889
T84	-.4085	-5.017	-.3494	-3.350	-.2933	-2.298
T85	-.8567	-10.110	-.7645	-7.039	-.7820	-5.845
T86	-1.2755	-14.937	-1.1804	-10.770	-1.2660	-9.402
T87	-1.6121	-18.728	-1.5201	-13.805	-1.6758	-12.368
T88	-2.0331	-22.412	-1.9813	-16.876	-1.9611	-14.177
R^2		.7416		.7003		.8220
No. Observations:		1265		917		348
Root MSE		.4166		.4181		.3796
Mean - LRPRICE		7.003		7.128		6.674

traditional critical value is 1.44, the Bayes large sample-adjusted critical value is 8.11, and the improvement in RMSE under the alternative hypothesis is

4.52%. Hence, although a tight criterion suggests rejection of the null, the more conservative Bayes and RMSE approaches lend only marginal support in favor of the alternative hypothesis.

Finally, we ran seven yearly regressions, one for each year from 1982 to 1988, and then compared the residual sums of squares from these regressions with that from the pooled model reported in column one of Table 5. Here we found more support for the notion of parameter instability. In particular, while the F-test statistic for the null hypothesis of parameter equality is 5.18 with a .01 critical value of 1.32, the Bayes critical value is larger at 9.03. However, the improvement in RMSE under the alternative hypothesis is substantial -- 15.76%. Thus, parameters do not appear to be stable. We will return to a discussion of parameter instability over time later in this section.

As discussed earlier, an alternative functional form specification involves including vintage and age dummies -- see equation (3), instead of the time and age dummies as in (2). Recall that regression results (including R^2, parameter estimates and standard errors) will vary somewhat when using the V-A specification rather than the T-A representation, and that this should not be surprising, for in (2) the total number of T-A-V dummy variable coefficients estimated directly is nine, while in (3) it is twelve. The summary statistics results found in the very bottom portion of Table 6 illustrate this variation among the various T-A and V-A specifications, for R^2 and RMSE values differ.[14]

One result of particular interest here concerns the age coefficients. As is seen in Table 6, with the V-A specification the age coefficients become negative and statistically significant, and increase in absolute value with age. We interpret these age coefficients, conditioning on vintage, as capturing the average decline in prices of surviving computer models given steady improvements

TABLE 6

PARAMETER ESTIMATES WITH TIME AND AGE, AND
WITH VINTAGE AND AGE DUMMY VARIABLE SPECIFICATIONS

	TIME AND AGE DUMMIES			VINTAGE AND AGE DUMMIES		
	POOLED	RETAIL	DISCOUNT	POOLED	RETAIL	DISCOUNT
AGE1	.1193 (.031)	.1296 (.037)	.0414 (.056)	-.2535 (.030)	-.2523 (.036)	-.2513 (.057)
AGE2	.1542 (.045)	.2352 (.059)	.0192 (.072)	-.5846 (.047)	-.5026 (.063)	-.6707 (.075)
AGE3	.2984 (.074)	.5333 (.112)	.1469 (.101)	-.8577 (.080)	-.5666 (.128)	-1.0561 (.112)
T83	-.3974 (.083)	-.2193 (.105)	-.8034 (.136)			
T84	-.4085 (.081)	-.3494 (.104)	-.2933 (.128)			
T85	-.8566 (.085)	-.7645 (.109)	-.7820 (.134)			
T86	-1.2755 (.085)	-1.1804 (.110)	-1.2660 (.135)			
T87	-1.6121 (.086)	-1.5201 (.110)	-1.6758 (.135)			
T88	-2.0331 (.091)	-1.9813 (.117)	-1.9611 (.138)			
V79				1.5830 (.183)	1.2007 (.254)	1.9415 (.271)
V80				1.0504 (.141)	.9474 (.227)	1.1670 (.205)
V81				.4454 (.148)	.5003 (.186)	.3439 (.243)
V83				.1646 (.093)	.0942 (.115)	.3536 (.156)
V84				-.1888 (.093)	-.2287 (.116)	-.0707 (.157)

V85				-.5502 (.096)	-.5869 (.120)	-.4144 (.164)
V86				-.9763 (.097)	-1.0051 (.123)	-.8583 (.162)
V87				-1.2928 (.098)	-1.3289 (.124)	-1.2157 (.161)
V88				-1.8130 (.107)	-1.8808 (.131)	-1.3605 (.205)
R^2	.7416	.7003	.8220	.7455	.7059	.8112
Root MSE	.4166	.4181	.3796	.4140	.4149	.3927

in new computers entering the market, i.e. as the average difference between the time and vintage effects. In a somewhat vague sense, therefore, these age coefficients capture the average effect of technical progress-induced obsolescence in our sample.

Since the interpretations and results from the T-A and V-A specifications differ considerably, and although our purpose of computing price indexes lends _a priori_ support to use of the T-A model specification in Table 5, one might still question whether using information from vintages in addition to that contained in the T-A model significantly improves model fit. In the previous section we noted that a full T-A-V specification is possible, provided that two variables are deleted from the V vector. Moreover, in our context a test for the null hypothesis that the α_v coefficients are simultaneously equal to zero can be interpreted as a specification test, providing information on whether the effects of unobserved and omitted characteristic variables are systematic among vintage comparisons, and/or whether equality constraints on characteristics parameters are invalid over vintages.

We therefore ran an additional regression in which eight vintage dummy variables were added to the model reported in column 1 of Table 5, and V82 and

V83 were deleted. The F-test statistic for the null hypothesis that $\alpha_v = 0$ is 5.94, the traditional .01 F-critical value is 2.51, the Bayes criterion value is 7.09, and the improvement in RMSE under the alternative hypothesis is 4.84%. Thus the traditional F-test suggests rather decisive rejection of the null hypothesis, the Bayes critical value of 7.09 is only slightly larger than the calculated F-statistic of 5.94, and the improvement in RMSE is almost up to our 5% threshhold. Hence, although the evidence is not clearcut, we interpret these results as providing some support for the alternative hypothesis, and therefore as admonishing us to assess our T-A specification in column 1 of Table 5 more closely, examining in particular what implicit parameter restrictions might be contributing to rejection of the null hypothesis.

This led us back to examine our earlier year-by-year regressions, and to look for patterns of parameter inequality over time. Our inspection revealed that although coefficients on a number of variables trended over time, the most marked trends were for coefficients on the LRAM, LMHZ, LHRDDSK and DOTHER variables. We then specified and estimated two additional models, one an overlapping sample model in which three separate regressions were run for the overlapping years 1982-84, 1984-86 and 1986-88, and the other a pooled 1982-88 run with four time-interaction variables added, LRAM*TC, LMHZ*TC, LHRDDSK*TC and DOTHER*TC, where TC is a time counter increasing annually from zero in 1982 to six in 1988. Results from these <u>overlapping</u> and <u>time-interaction</u> regressions are presented in Table 7.

The results presented in Table 7 represent an improvement in the model specification, accounting somewhat for the considerable variation among parameter estimates over time. For example, in the 1982-84, 1984-86 and 1986-88 regressions, coefficient estimates on LRAM, LHRDDSK, LNUMFLP and DOTHER fall

Table 7

Final Set of Regression Runs
Pooled 1982-88 with Time Interactions, and
Separate 1982-84, 1984-86, 1986-88 Model Runs

	Overlapping Regressions						1982-88 Model with Time Interactions	
	1982-84 Model		1984-86 Model		1986-88 Model			
Variable	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Intercept	4.1805	19.447	4.6522	22.881	4.5653	29.790	3.7782	25.066
LRAM	.4622	9.872	.1925	4.768	.1652	6.883	.6297	15.857
LRAM*TC							-.0855	-9.416
LMHZ	.0818	1.047	.4041	6.521	.4580	9.427	.1968	2.846
LMHZ*TC							.0370	2.228
LHRDDSK	.2405	7.591	.2090	12.190	.1603	20.061	.2302	7.612
LHRDDSK*TC							-.0137	-2.301
LNUMFLP	.6089	5.880	.3916	4.976	.1625	2.753	.3271	6.644
LSLOTS	.2429	5.453	.2613	8.136	.1134	4.930	.1556	7.965
AGE1	.1527	2.030	.1321	2.725	.1593	5.134	.1410	4.835
AGE2	.0217	0.172	.0793	0.983	.1701	3.841	.1593	3.733
AGE3	.3827	2.644	.1758	1.070	.1907	2.342	.2496	3.525
PROC16	.1429	1.751	.1255	2.302	.2824	7.338	.2170	6.315
PROC32			.2736	1.097	.6040	9.392	.6152	9.573
DBW	.1538	2.163	.0644	1.387	-.1190	-3.771	.0013	0.046
DCOLOR	.3498	1.547	-.0070	-0.042	-.0249	-0.340	.0590	0.831
DPORT	.0770	0.890	.4723	7.067	.5019	11.217	.3967	10.365
DEXTRA	.0283	0.166	.2670	2.201	.3137	3.092	.2132	2.615
DDISC	-.3445	-5.196	-.2430	-4.778	-.3053	-9.508	-.2946	-10.061
DAPPLE	.2993	1.925	.4641	4.158	.2439	2.911	.2199	3.047
DCMDRE	-.4662	-2.331	-.3954	-2.681	-.3076	-3.292	-.3672	-4.409
DCMPQ	.4631	2.535	.2757	2.756	.0913	1.395	.1929	3.109
DNEC	.2916	1.686	-.0582	-0.482	-.0580	-0.770	.0399	0.578
DRDIOSH	.4379	3.113	-.0387	-0.335	-.3162	-3.964	.0704	1.056
DPCLIM			-.3583	-1.998	-.5025	-5.331	-.5136	-5.247
DOTHER	.2680	2.408	.1467	1.996	-.1316	-2.878	.2607	3.498
DOTHER*TC							-.0648	-4.405
T83	-.5203	-5.631					-.2552	-3.144
T84	-.6203	-6.173					.0029	0.034
T85			-.4015	-7.196			-.0787	-0.710
T86			-.7694	-12.584			-.0319	-0.218
T87					-.3365	-9.915	.1372	0.722
T88					-.7561	-18.667	.2680	1.110
R^2		.8310		.7336		.7810		.7668
No. Observations:		247		428		904		1265
Root MSE		.4183		.3889		.3595		.3965

continuously, while that on LMHZ increases. Trends are also apparent in several other coefficients. Moreover, when the pooled 1982-88 regression model with time-interactions is estimated, negative and statistically significant estimates are obtained for LRAM*TC (-.0885), LHRDDSK*TC (-.0137) and DOTHER*TC (-.0648), while that on LMHZ*TC (.037) is positive and significant. Hence both of these more general specifications appear to provide improved estimates.

To check further on the validity of these two specifications, we added to each regression the set of eight dummy vintage variables, and then tested the null hypothesis that $\alpha_V = 0$. Our results are more satisfying, and lend qualified support for the models reported in Table 7. In particular, as shown in Table 8 below, for 1982-84 and 1984-86, the calculated F-statistics are less than the .01 critical values, for 1986-88 the calculated F-statistic is larger, but in all three cases F-statistics are smaller than the Bayes criterion, and the improvement in RMSE with vintage variables included is less than 1.5%. Hence, for all three overlapping models, whatever the effects of omitted and unobserved characteristics, they do not appear to be systematic among vintage comparisons.

With the pooled 1982-88 time-interaction model, results are roughly similar to those from the overlapping models. The calculated F-statistic is larger than the .01 critical value but less than half the Bayes criterion, and the improvement in the RMSE when vintage variables are added is less than 1%. Thus there is little basis to choose among these two specifications. However, we expect that the constant change in parameters implied by the interactive time counter would become increasingly inappropriate as additional time observations were added. On this criterion, therefore, we have a mild preference for the specification involving three overlapping regressions.

Table 8

Test Results for Null Hypothesis that Vintage Effects are Zero
Separate 1982-84, 1984-86, 1986-88 Models, and
Pooled 1982-88 Model with Time Interactions

Model	Calculated F-statistic	.01 Critical F-value	Bayes-Leamer Criterion	% Change in RMSE
Overlapping 1982-84	2.72	3.32	5.23	1.45%
Overlapping 1984-86	3.30	3.78	6.20	0.84%
Overlapping 1986-88	5.85	3.32	6.72	1.08%
Time Interaction 1982-88	3.55	2.51	7.07	0.82%

Although further experimentation with other combinations of characteristics variables would most likely be useful, we now move on to using several of the above most promising specifications to construct quality-adjusted price indexes for PC's.

V. PRICE INDEXES

Price indexes can be constructed in a variety of ways using the results of our various hedonic price equations. Although possibilities are limited when quantity sales data on the various models are unavailable, numerous procedures can be implemented given enough available data. In this section we construct and comment on several price indexes, all based on our hedonic regression equations, but varying in their interpretation and in their use of parameter estimates and quantity weights.

We begin with price indexes based on direct transformation of estimated hedonic price coefficients, and interpreted as price indexes holding quality constant over time. In the first three rows of Table 9 we present implicit PC price indexes computed directly from the three T-A regression equations reported in Table 5, constructed here simply as the exponentiated estimated coefficients on the time dummy variables, with T_{82} set to zero. The values in parentheses are percent changes from the previous year, computed as $100*(PI_t - PI_{t-1})/PI_{t-1}$, where PI is the price index. Note that the price indexes in the T-A model are the estimated time effects from regressions holding age and other characteristics constant. Overall, we see that average annual growth rates (AAGR) are similar for the pooled, retail and discount equations (about -28% per year), although the deflators for discounted models tend to be somewhat unstable from 1983 to 1985.

In the second set of three rows in Table 9, implicit price indexes are presented based on direct exponentiation of the estimated vintage coefficients from the V-A specifications in Table 6. The interpretation of these price indexes is slightly different -- they are not time effects, but rather are based on vintage coefficients, holding age and other characteristics fixed. As is seen in Table 9, these price indexes suggest slightly slower declines in quality adjusted prices than those based on T-A regressions (especially for discount models), and also reveal greater instability, particularly between 1982 and 1985.

One might think of these V-A price indexes as tracing out quality-adjusted price deflators for various vintages having AGE = 0 (since price indexes for 1979 to 1988 are computed directly from the V79 to V88 vintage coefficients, assuming AGE = 0), but estimation of the underlying coefficients is based on a sample including models of all ages. An alternative procedure for constructing a price index for new models only -- an index that might be construed as tracing out the technological "frontier" -- is to estimate parameters from a data sample restricted to new models, i.e. to models with AGE = 0. Implicit price indexes computed from such a regression are reported in the row named "New Models Only" in Table 9.[15] There it is seen that a "new model only" price index declines more rapidly than those based on full-sample T-A and V-A specifications; in particular, the AAGR from 1982 to 1988 is -30.15%.

The two final implicit price indexes computed directly from hedonic regression equations without use of quantity sales weights are given in Table 9 in the rows marked "Overlapping" and "Time Interactions". The overlapping price indexes are based on the three overlapping regressions reported in Table 7. They are computed by directly exponentiating the coefficient estimates on the time dummy variables, linked so that, for example, the implicit 1985 and 1986 price indexes are the products of the exponentiated coefficients for 1984 and 1985, and 1984 and 1986, respectively. Notice that with an overlapping index procedure, the quality weights are constant only for sub-periods, and coefficient estimates reflect varying sample means among sub-periods. Interestingly, the overlapping price indexes fall at almost the same AAGR as that based on a "new models only" regression, although the overlapping price indexes fall more rapidly in the earlier years.

The computation of price indexes based on the time-interaction model requires use of sample characteristics data, not just values of estimated coefficients. For example, using parameter estimates on the time interaction terms reported in Table 7, for the log-change in quality-adjusted prices between year t and year t-1 we first compute

$$\ln \tilde{p}_t - \ln \tilde{p}_{t-1} = (\alpha_t - \alpha_{t-1}) - .0855*LRAM'_t + .0370*LMHZ'_t - .0137*LHRDDSK'_t - .0648*TC_t*DOTHER'_t, \qquad (7)$$

where the $'_t$ on LRAM, LMHZ, LHRDDSK and DOTHER refers to the sample mean of these variables between years t and t-1. To calculate the price index, we simply cumulate the values in (7) over 1982-1988 (letting $\alpha_{1982} = 0$), and then exponentiate them. Results from such a computation are presented in the "Time

TABLE 9

ALTERNATIVE IMPLICIT REAL QUALITY-ADJUSTED PRICE INDEXES FOR PC'S
Based on Direct Hedonic Regression Estimates
(Percent Negative Change from Previous Year in Parentheses)

Procedure	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	82-88 AAGR
T-A Pooled				1.000	.672	.665	.425	.279	.200	.131	
					(33%)	(1%)	(36%)	(34%)	(28%)	(35%)	-28.7%
T-A Retail				1.000	.803	.705	.466	.307	.219	.138	
					(20%)	(12%)	(34%)	(34%)	(29%)	(37%)	-28.1%
T-A Discount				1.000	.448	.746	.458	.282	.187	.141	
					(55%)	(-67%)	(39%)	(38%)	(34%)	(25%)	-27.9%
V-A Pooled	4.869	2.859	1.561	1.000	1.179	.828	.577	.377	.274	.163	
		(41%)	(45%)	(36%)	(-18%)	(30%)	(30%)	(35%)	(27%)	(41%)	-26.1%
V-A Retail	3.322	2.579	1.649	1.000	1.099	.796	.556	.366	.265	.152	
		(29%)	(36%)	(39%)	(-10%)	(28%)	(43%)	(34%)	(28%)	(43%)	-26.9%
V-A Discount	6.969	3.212	1.410	1.000	1.424	.932	.661	.424	.296	.256	
		(54%)	(56%)	(29%)	(-42%)	(35%)	(29%)	(36%)	(43%)	(14%)	-20.3%
New Models Only				1.000	.716	.620	.420	.266	.195	.116	
					(28%)	(13%)	(32%)	(37%)	(27%)	(41%)	-30.2%
Overlapping				1.000	.594	.538	.360	.249	.178	.117	
					(41%)	(9%)	(33%)	(31%)	(29%)	(34%)	-30.1%
Time Interactions				1.000	.560	.494	.296	.182	.129	.086	
					(44%)	(12%)	(40%)	(39%)	(29%)	(33%)	-33.6%

Note: The values in parentheses are percent changes from the previous year, computed as $100*(PI_t - PI_{t-1})/PI_{t-1}$, where PI is the price index.

Interactions" row of Table 9. There we see that use of this price index procedure results in a very substantial rate of price decline -- more than 33% per year. This large decline reflects the fact that sample means of the variables are all increasing with time, and these means are multiplied by the relatively large negative coefficients in (7).

One important problem with each of the above price indexes is that they fail to reflect changes over time in the mix of models. Recall that the direct hedonic regression coefficients in the T-A models can be interpreted as holding quality constant either by fixing the base of characteristic values over time, or by fixing their valuation (parameter estimates). In a world with rapidly evolving new technologies, the notion of a fixed characteristic base as portraying representative transactions becomes increasingly inappropriate. What would be preferable is an index number procedure that accounts for compositional changes in models over time.[16] Such a computation requires, of course, quantity and revenue sales data by model by year. As our final index number computations, we now consider a Divisia index that weights quality-adjusted prices of models by their revenue shares.

Specifically, our calculation of a (Tornqvist approximation to the) Divisia index proceeds as follows. First consider a model j observed in both time periods 0 and 1. Let

$$\ln P_{1,j} = Z'_{1,j}\hat{\beta} + \hat{\alpha}_1 + \epsilon_{1j}, \text{ and} \tag{8a}$$

$$\ln P_{0,j} = Z'_{0,j}\hat{\beta} + \hat{\alpha}_0 + \epsilon_{0j}, \tag{8b}$$

where $Z_{1,j}$ and $Z_{0,j}$ are vectors of all regression variables except for the time dummy variables in years 1 and 0, the $\hat{\alpha}$'s are estimated coefficients on the time dummy variables, and the ϵ's are least squares residuals. This implies that

$$\ln P_{1,j} - \ln P_{0,j} = (Z'_{1,j} - Z'_{0,j})\hat{\beta} + (\hat{\alpha}_1 - \hat{\alpha}_0) + (\epsilon_{1j} - \epsilon_{0j}),$$

which can be rearranged to yield the expression

$$\ln P_{1,j} - \ln P_{0,j} - (Z'_{1,j} - Z'_{0,j})\hat{\beta} = (\hat{\alpha}_1 - \hat{\alpha}_0) + (\epsilon_{1j} - \epsilon_{0j}). \tag{9}$$

The left-hand side of (9) states that the log-change in the quality-adjusted price of model j from 0 to 1 equals the change in observed prices minus the change in quality, where quality is evaluated using least squares regression coefficients and values of the characteristics. Alternatively, the right-hand side of (9) states that an equivalent way of computing the log-change in the quality-adjusted price of model j is simply to sum the difference in estimated time dummies plus the difference in least squares residuals. The choice of which of these two methods to employ in computing quality-adjusted prices can be based simply on relative computational convenience.

Several other features of (9) are worth noting. First, if there is no change in the characteristics of model j between 0 and 1 (i.e., the model has not become a new version in period 1), then $(Z'_{1,j} - Z'_{0,j}) = 0$, this term drop outs of (9), and in essence the quality-adjusted log price change is computed using the traditional matched model procedure. Second, if the least squares residual is the same in the two time periods, i.e. if $\epsilon_{1j} - \epsilon_{0j} = 0$, then the log-change in quality-adjusted prices is simply equal to the change in the time dummy coefficients. Note that residuals have a useful interpretation in the hedonic price equation, for they provide evidence on whether relative to the overall market, a particular model is over- or under-priced.[17] An interesting issue concerns the relationship between these residuals and the revenue shares garnered by each model. Since for each year the sum of residuals is zero, we would expect that if shares are uncorrelated with residuals, it would also be the case that $\sum s_j \epsilon_j \approx 0$.

Once (9) is computed for every model j in years 0 and 1, the log-change in quality adjusted prices over all models is calculated as the revenue share-weighted sum of the individual model j log-changes in quality-adjusted prices,

$$\ln \tilde{P}_1 - \ln \tilde{P}_0 = \sum_{j=1}^{J} \bar{s}_j \,(\ln \tilde{P}_{1,j} - \ln \tilde{P}_{0,j}) = \sum_{j=1}^{J} \bar{s}_j (\epsilon_{1j} - \epsilon_{0j}) + \hat{\alpha}_1 - \hat{\alpha}_0 \quad (10)$$

where the ˜ superscript is the quality-adjusted price (computed for individual j models using either side of equation (9)), $\bar{s}_j$ is the arithmetic mean of $s_{j,1}$ and $s_{j,0}$, and s_j is the share of model j's value of shipments in the total value of shipments over all models in the appropriate time period.

The calculation in (9) is feasible only when model j is part of a surviving cohort of models. In fact, however, some models exit the market each year, while others enter. To account for these entering and exiting models, several adjustments must be made to (9) and (10).

Consider the case of a model that enters the market in time period 1. Obviously, its price cannot be observed in period 0, and thus use of (9) to compute a quality-adjusted price index is not feasible. One can, however, use the estimated hedonic regression equation to predict such missing prices. In particular, substitute the right-hand side of (8b) into (9), let the values of Z_{1j} and Z_{0j} be equal, set ϵ_0 to zero, and rearrange. This yields an expression for quality-adjusted log-price changes for entering models in period 1, computed in two alternative but equivalent ways as

$$\ln P_{1,j} - Z_{1,j}'\hat{\beta} - \hat{\alpha}_0 = \hat{\alpha}_1 - \hat{\alpha}_0 + \epsilon_{1j}. \tag{11}$$

Similarly, for exiting models that were observed in period 0 but not in period 1, we employ hedonic regression procedures to predict the price of that model would it have survived to period 1. Specifically, we substitute (8a) into (9), set $Z_{1j} = Z_{0j}$ and $\epsilon_1 = 0$, and then rearrange. This yields the appropriate log-change in the quality-adjusted prices for exiting models as

$$Z_{0,j}'\hat{\beta} + \hat{\alpha}_1 - \ln P_{0,j} = \hat{\alpha}_1 - \hat{\alpha}_0 - \epsilon_{0j}. \tag{12}$$

Once these log-changes in quality-adjusted prices are computed for all continuing, entering and exiting models, we calculate revenue shares (setting $s_{0,j}$ to zero for entering models, and $s_{1,j}$ to zero for exiting models, thereby effectively using one-half of the last or first observed share weight), and then compute an aggregate log-change in quality-adjusted prices over all models using (10).

Several other points are worth noting. First, an interesting feature of (9), (11) and (12) is that they employ as information the values of the least squares residuals. Hence the Divisia quality-adjusted index number procedure takes into account whether those models that exited (or entered) had prices above or below the average quality-adjusted prices. Note, however, that the weight given these exiting and entering models is likely to be minor, since their average revenue share in periods 0 and 1 is in most cases rather small.

Second, empirical implementation of this Divisia index number procedure requires data on value of shipments by model. Proprietary data on shipments, installations and value of shipments by model and year for about 950 of the 1265 models in our estimation sample were kindly provided us by the International Data Corporation. These data formed the basis of our share-weights used in (10).

Restricting our sample to models covered by the IDC data set and computing revenue values by model-year as the product of the IDC estimates of average price paid and number of models shipped, we have computed Divisia quality-adjusted price indexes separately for continuing, entering and exiting models, and for selected aggregates.[18] Our results, using parameter estimates from the T-A pooled and T-A overlapping estimated models, are presented in Table 10. A number of results are worth noting.

We begin with results from the T-A pooled estimation. Inspection of the top row of Table 10 reveals that the real quality-adjusted Divisia price index for all PC models declined at an AAGR of -28.2% from 1982 to 1988, virtually identical to the -28.7% AAGR of the direct hedonic T-A pooled price index. A striking difference between these two price indexes, however, is that the Divisia index reveals a much smoother decline over time, with year-to-year

TABLE 10

ALTERNATIVE DIVISIA REAL QUALITY-ADJUSTED PRICE INDEXES FOR PC'S
Based on T-A Pooled and T-A Overlapping Regression Estimates
(Percent Negative Change from Previous Year in Parentheses)

Regression and Sample	1982	1983	1984	1985	1986	1987	1988	82-88 AAGR
T-A Pooled Estimation:								
All Computer Models	1.000	.638	.510	.385	.283	.188	.136	
		(36%)	(20%)	(25%)	(26%)	(34%)	(28%)	-28.2%
Continuing Models Only	1.000	.580	.438	.330	.247	.160	.112	
		(42%)	(24%)	(25%)	(25%)	(35%)	(30%)	-30.6%
Entering Models Only	1.000	.716	.562	.379	.270	.201	.179	
		(28%)	(22%)	(33%)	(29%)	(26%)	(11%)	-24.9%
Exiting Models Only	1.000	.804	1.188	1.005	.682	.410	.263	
		(20%)	(-48%)	(15%)	(32%)	(40%)	(36%)	-20.0%
T-A Overlapping Estimation:								
All Computer Models	1.000	.576	.465	.359	.282	.193	.140	
		(42%)	(19%)	(23%)	(21%)	(32%)	(27%)	-28.0%
Continuing Models Only	1.000	.542	.422	.317	.250	.160	.113	
		(46%)	(22%)	(25%)	(21%)	(36%)	(29%)	-30.5%
Entering Models Only	1.000	.756	.586	.453	.355	.261	.218	
		(24%)	(22%)	(23%)	(22%)	(26%)	(16%)	-22.4%
Exiting Models Only	1.000	.591	.780	.638	.440	.299	.197	
		(41%)	(-32%)	(18%)	(31%)	(32%)	(34%)	-23.7%

Note: The values in parentheses are percent changes from the previous year, computed as $100*(PI_t - PI_{t-1})/PI_{t-1}$, where PI is the price index.

declines ranging between 20% and 37%, whereas year-to-year declines for the direct T-A pooled index vary from 1% to 36%.

Second, changes in the price indexes for the continuing models are quite different from those models entering and exiting between 1982 and 1988. Specifically, if one examines the next three rows in Table 10, one sees that the price declines of the continuing models were on average larger (-30.6%)

than those for the entering (-24.9%) and exiting (-20.0%) models. An implication is that vendors of incumbent models responded quite aggressively to the competition provided by entrants, dropping prices at relatively high rates. Note that these results can be reconciled with the econometric findings reported earlier that in the T-A specification, estimated coefficients on the age variables were positive, provided one interprets the latter result as reflecting selectivity due to, perhaps, unobserved positive quality differentials among the survivors as vintages progress.

Third, the pricing strategies employed for entering and exiting models are quite different. Over the 1982-88 time period, quality-adjusted price declines for entering models were on average larger (-24.9%) than those for exiting models (-20.0%). However, while the quality-adjusted prices were on average flat between 1982 and 1985 for exiting models, these models exhibited very large price declines from 1985 to 1988.

In the final set of four rows of Table 10, we present Divisia price indexes using parameter estimates from the 1982-84, 1984-86 and 1986-88 overlapping regressions. The remarkable result here is that in spite of using a rather different set of regressions, average annual growth rates of quality-adjusted prices from 1982 to 1988 are hardly affected. Specifically, the AAGR for all computer models based on the pooled and overlapping regressions are, respecitvely, -28.2 and -28.0%; for continuing models, -30.6% and -30.5%; for entering models, -24.9% and -22.4%; and for exiting models, -20.0% and -23.7%. Although year-to-year variations occur between the Divisia pooled and overlapping regression price indexes, AAGR are reasonably robust.

VI. SUMMARY REMARKS

The simultaneous existence of incumbent, entering and exiting models raises issues of product heterogeneity in the microcomputer market, the nature of price and quality competition, and creates ambiguity in how one constructs and interprets price indexes. Those are the issues on which we have focused in this paper.

Specifically, we have reported results from estimation of a variety of hedonic regression equations using an unbalanced panel data set for 1265 model-years from 1982 to 1988, and have developed and implemented empirically a specification test for selecting preferable hedonic price equations. We have discussed in detail the alternative interpretation of dummy variable coefficients in models having time and age, vintage and age, and all of the time, age and vintage dummy variables as regressors. Based on these estimated hedonic price equations, we then computed quality-adjusted price indexes using a variety of procedures. This provided us with indexes having varying interpretations -- constant average quality price indexes, price indexes for new models only, and quality-adjusted price indexes portraying representative transactions that take into account the changing model composition in our sample over time. Not surprisingly, average annual growth rates for these varying price indexes also differed, although all showed a substantial decline in quality-adjusted prices over the 1982-88 time period.

Our research in this paper is preliminary, and much still remains to be done. One item high on our research agenda involves obtaining model-specific performance measures for specific numerical tasks, such as the number of instructions executed per unit of time, and then re-doing our hedonic regressions with such performance measures added as regressors. Moreover, issues of parameter instability and choice of variables to include in the set of characteristics are also potentially important, and need further examination. Finally, our least squares residuals provide economic information either on over- or under-pricing of models relative to the market as a whole, or else reflect the effects of omitted variables or differential market power in different PC market segments. Hence an interesting application involves examining in greater detail the relationships among residuals, entering, continuing, and exiting models, and realized market shares.

FOOTNOTES

[1]See, for example, the classic study by Gregory Chow [1957], as well as more recent ones by Robert B. Archibald and William S. Reece [1978], Robert J. Gordon [1989,1990], Robert Michaels [1979], Stephen D. Oliner [1986], and Jack E. Triplett [1989a].

[2]A very brief discussion of PC's is presented in Robert J. Gordon [1989,1990]. Also see the unpublished paper by Brian Catron [1989].

[3]Hedonic regression methods and their interpretation are discussed in, among others, Zvi Griliches [1961,1971,1988], Jack E. Triplett [1986], and Ernst R. Berndt [1990, chapter 4]. Theoretical foundations for interpreting hedonic price equations are found in, among others, Sherwin Rosen [1974] and Dennis Epple [1987]. For an historical discussion on the incorporation of hedonic regression methods into official price indexes, see Triplett [1990].

[4]Precisely how the BEA PC price index is constructed is not clear. According to David W. Cartwright and Scott D. Smith [1988, p. 22], "For personal computers (PC's), a matched model index was introduced in 1987. It is now constructed using price changes of IBM PC's, judgmentally adjusted by BEA to reflect price changes for other models, for 1983 and price changes of models sold by IBM and three additional manufacturers for 1984-87."

[5]The first PC advertising appeared in the New York Times in 1981.

[6]A model is defined as discounted if it is sold by a vendor other than the brand-name manufacturer. Thus, for example, IBM models sold by IBM are considered as retail price observations, while IBM models sold by Computerland or 47th Street Photo are treated as discounted price observations. PC Limited models (sold only by mail order from PC Limited) are treated as retail, since PC Limited is the only vendor.

[7]However, one cannot identify parameters in a full quadratic expansion of the three variables, due to the identity in (1). For a discussion in the context of age, period and cohort models, see Stephen Fienberg and William Mason [1985].

[8]A related discussion of this issue in the context of age, period and cohort effects in earnings equations is presented by James Heckman and Richard Robb [1985].

[9]See, for example, David W. Cartwright and Scott D. Smith [1986], Rosanne Cole et al. [1986], Robert J. Gordon [1989] and Jack E. Triplett [1989a].

[10]See especially Robert Hall [1971], p. 248.

[11]There is intuitive appeal to this additional normalization. Hall defined the price index as the product of vintage effects (embodied technical progress), depreciation, and time (disembodied technical progress). Thus the logarithm of the price index is the sum of these three effects, each in rates of growth. To normalize the level of the price index, one normalizes levels of each of the three effects, i.e., one deletes one variable from each of the T, V and A dummy variable sets, and normalizes relative to that variable. But in addition, one must normalize at least one of the growth rates, since the product of the three effects implies that components are unidentified. This additional normalization is accomplished by deleting an additional vintage

variable, thereby yielding a contrast in levels of the logarithmic regression, which is equivalent to a normalization in growth rates of one of the three components. For additional discussion, see Robert Hall [1971].

[12]It is worth noting here that the choice of which two dummy variables to delete from the V vector is arbitrary in the sense that goodness of fit and numerical values of least squares estimates of α and the β's will be unaffected. However, the interpretation and numerical values of the least squares estimates of the α_t, α_a and α_v will depend on this choice.

[13]Leamer has derived this to equal $(n-k)\cdot(n^{q/n} - 1)/q$, where n is sample size, k is the number of free parameters estimated in the unconstrained regression, and q is the number of parameter restrictions. For an application of Leamer's adjustment to the standard F-test procedure in the context of large samples, see Makoto Ohta and Zvi Griliches [1976].

[14]While not reported here for reasons of space, it is worth noting that the slope coefficient estimates differ between the T-A and V-A specifications, although in many cases the differences are not large.

[15]The underlying regression equation is of the same form as in Table 5, except that age variables are deleted. Price indexes are computed directly by exponentiating the estimated coefficients on the time dummy variables.

[16]For a recent discussion of weighting issues in the context of compositional changes, see Jack A. Triplett [1989b] and the references cited therein.

[17]This under- or over-pricing might also reflect of course the effects of unobserved omitted variables, or of differential market power in differing segments of the PC market.

[18]We divided revenues among retail and discount listings of the same model in proportion to the relative number of listings. It is also worth noting that mean values of the revenue shares of continuing, entering and exiting models from 1982 to 1988 are 54%, 26% and 20%, respectively. There is considerable variation in these shares over our sample time period, however.

REFERENCES

Archibald, Robert B. and William S. Reece [1978], "Partial Subindexes of Input Prices: The Case of Computer Services," Southern Economic Journal, 46:2, October, 528-540.

Arrow, Kenneth J. [1960], "Decision Theory and the Choice of a Level of Significance for the t-Test," in I. Olkin, W. Hoeffding, S. G. Gurye, W. G. Madow and H. B. Mann, eds., Contributions to Probability and Statistics, Stanford: Stanford University Press, 70-78.

Berndt, Ernst R. [1990], "The Measurement of Quality Change: Constructing an Hedonic Price Index for Computers using Multiple Regression Methods," ch. 4 in Berndt, The Practice of Econometrics: Classic and Contemporary, Reading, MA: Addison-Wesley Press, forthcoming.

Cartwright, David W. and Scott D. Smith [1988], "Deflators for Purchases of Computers in GNP: Revised and Extended Estimates, 1983-1988," Survey of Current Business, 68:4, April, 22-23.

Cartwright, David W. [1986], "Improved Deflation of Purchases of Computers," Survey of Current Business, 66:3, March, 7-9.

Catron, Brian [1989], "Price Measurement for Computer Hardware: A Demonstration of Quality Adjustment Techniques," US Department of Labor, Bureau of Labor Statistics, Office of Prices and Living Conditions, Working Paper, April 14.

Chow, Gregory C. [1967], "Technological Change and the Demand for Computers," American Economic Review, 57:5, December, 1117-1130.

Cohen, Jeremy M. [1988], "Rapid Change in the Personal Computer Market: A Quality-Adjusted Hedonic Price Index, 1976-1987," unpublished S.M. thesis, Massachusetts Institute of Technology, Alfred P. Sloan School of Management, May.

Cole, Rosanne, Y. C. Chen, Joan A. Barquin-Stolleman, Ellen Dulberger, Nurhan Helvacian and James H. Hodge [1986], "Quality-Adjusted Price Indexes for Computer Processors and Selected Peripheral Equipment," Survey of Current Business, 66:1, January, 41-50.

Dulberger, Ellen R. [1989], "The Application of an Hedonic Model to a Quality Adjusted Price Index for Computer Processors", in Dale W. Jorgenson and Ralph Landau, eds., Technology and Capital Formation, Cambridge, MA: The MIT Press, 37-75.

Epple, Dennis [1987], "Hedonic Prices and Implicit Markets: Estimating Demand and Supply Functions for Differentiated Products," Journal of Political Economy, 95:1, January, 59-80.

Fienberg, Stephen E. and William M. Mason [1985], "Specification and Implementation of Age, Period and Cohort Models," in William M. Mason and Stephen E. Fienberg, eds., Cohort Analysis in Social Research, New York: Springer-Verlag, 45-88.

Fisher, Franklin M., John J. McGowan and Joen E. Greenwood [1983], Folded, Spindled and Mutilated: Economic Analysis and U.S. v. IBM, Cambridge, MA: The MIT Press.

Gordon, Robert J. [1990], The Measurement of Durable Goods Prices, Chicago: Univesity of Chicago Press for the National Bureau of Economic Research.

Gordon, Robert J. [1989], "The Postwar Evolution of Computer Prices," chapter 3 in Dale W. Jorgenson and Ralph Landau, eds., Technology and Capital Formation, Cambridge, MA: MIT Press, 77-125.

Griliches, Zvi [1988], "Postscript on Hedonics," in Zvi Griliches, Technology, Education, and Productivity, New York: Basil Blackwell, Inc., 119-122.

Griliches, Zvi [1971], "Introduction: Hedonic Prices Revisited," in Zvi Griliches, ed., Price Indexes and Quality Change: Studies in New Methods of Measurement, Cambridge, MA: Harvard University Press, 3-15.

Griliches, Zvi [1961], "Hedonic Price Indexes for Automobiles: An Econometric Analysis of Quality Change," in The Price Statistics of the Federal Government, General Series No. 73, New York: Columbia University Press for the National Bureau of Economic Research, 137-196. Reprinted in Zvi Griliches, ed., Price Indexes and Quality Change: Studies in New Methods of Measurement, Cambridge, MA: Harvard University Press, 1971, 55-87.

Hall, Robert E. [1971], "The Measurement of Quality Change from Vintage Price Data," chapter 8 in Zvi Griliches, ed., Price Indexes and Quality Change, Cambridge, MA: Harvard University Press, 240-271.

Heckman, James J. and Richard Robb [1985], "Using Longitudinal Data to Model Age, Period and Cohort Effects in Earnings Equations," in William M. Mason and Stephen E. Fienberg, eds., Cohort Analysis in Social Research, New York: Springer-Verlag, 137-150.

Kim, Amy Y. [1989], "Hedonic Price Indices and An Examination of the Personal Computer Market," Harvard College, honors undergraduate thesis, Department of Economics, March.

Leamer, Edward E. [1978], Specification Searches: Ad Hoc Inference with Nonexperimental Data, New York: John Wiley & Sons.

Michaels, Robert [1979], "Hedonic Prices and the Structure of the Digital Computer Industry," Journal of Industrial Economics, 27:1, March, 263-275.

Ohta, Makoto and Zvi Griliches [1976], "Automobile Prices Revisited: Extensions of the Hedonic Hypothesis," in Nestor E. Terleckyj, ed., Household Production and Consumption, Studies in Income and Wealth, Vol. 40, New York: Columbia University Press for the NBER, 325-390.

Oliner, Stephen D. [1986], "Depreciation and Deflation of IBM Mainframe Computers: 1980-1985," Washington, DC: Federal Reserve Board, unpublished manuscript, June.

Rosen, Sherwin [1974], "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," Journal of Political Economy, 82:1, January/February, 34-55.

Triplett, Jack E. [1990], "Hedonic Methods in Statistical Agency Environments: An Intellectual Biopsy," Washignton, DC: Bureau of Economic Analysis, Discussion Paper 36, Revised January 1990. Forthcoming in Ernst R. Berndt and Jack E. Triplett, eds., Fifty Years of Economic Measurement, Chicago: University of Chicago Press for the National Bureau of Economic Research.

Triplett, Jack E. [1989a], "Price and Technological Change in a Capital Good: A Survey of Research on Computers," chapter 4 in Dale W. Jorgenson and Ralph Landau, eds., Technology and Capital Formation, Cambridge, MA: MIT Press, 127-213.

Triplett, Jack E. [1989b], "Superlative and Quasi-Superlative Indexes of Price and Output for Investment Goods: Office, Computing and Accounting Machinery," Washington, DC: U. S. Department of Commerce, Bureau of Economic Analysis, Discussion Paper 40, Revised July.

Triplett, Jack E. [1986], "The Economic Interpretation of Hedonic Methods," Survey of Current Business, 86:1, January, 36-40.