A HEDONIC STUDY OF PREPACKAGED SOFTWARE

by

Robert John McCahill

Thesis submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

MASTER OF ARTS

in

Economics

APPROVED:

_____
Nancy Wentzler
Committee Chairperson

_____
William R. Porter

_____
Brian K. Reid

May 22, 1997
Falls Church, Virginia

Key Words: Hedonics, Prices, Software, Spreadsheet, Word Processing, Characteristics

A HEDONIC STUDY OF PREPACKAGED SOFTWARE

by

Robert John McCahill

Nancy Wentzler, Committee Chairperson

Economics

ABSTRACT

This study verifies previous econometric research which found that spreadsheet software prices, when adjusted for quality improvement, decline over the period 1986-1993. New econometric work is presented for prepackaged word processing software. Using objective criteria for variable selection, the model yields declining quality-adjusted prices for word processing software over the period 1985-1994.

Acknowledgments

The author wishes to thank research and thesis committee professors Nancy Wentzler, Russ Porter, and Brian Reid for their insightful comments and suggestions. Discussions with Stephen Oliner and Daniel Sichel of the Federal Reserve Board were extremely helpful in the early stages of this project. Professor David Meiselman, director emeritus, and department secretary Aniene Porter, deserve special recognition for their continued encouragement and support.

Contents

1 Purpose of this Study

This study takes an econometric approach to determine price movements for prepackaged software, adjusting for changes in quality. This research will attempt to confirm results of earlier hedonic studies on spreadsheet prepackaged software and will introduce a new hedonic function for word processing prepackaged software.

2 Controlling for the Effects of Quality Change on Prices

In order to make informed decisions on the real direction of the economy or the firm, policy makers need a tool to adjust prices for inflation. The "matched-model" technique, typically used by the BLS to develop the Producer Price Indexes (PPI), matches essentially identical items in adjacent periods, leaving new products out of the sample. Because items have the same characteristics, any change in price is expected to be a pure change, unrelated to quality differences.

This technique may be flawed for two reasons. First, the models included in the sample may not be representative of the entire sample. For instance, if there are a substantial number of new products on the market, which have no base period to compare to and which by definition therefore are excluded from the model (until it's second period of sale at the earliest) the sample may not be representative. Likewise, introduction of improved technology can lower the bidding price of old-technology products to the point they cannot compete, and simply drop from production. In this case the base period observation is excluded for lack of a current period match. Second, those models included may indeed not be strictly identical, but close enough in features that they are included. The alternative of dropping the observation is outweighed by the need to have an adequate number of models to compare. See Triplett (1986) for a discussion of inherent matched-model weaknesses for products having high rates of quality change. The BLS at present has no price index for prepackaged software.

For rapidly evolving products which typically appear and disappear with maddening frequency, hedonic techniques may be superior to traditional matched-model techniques: by either estimating prices of missing models, or estimating the entire price index from regression coefficients. Hedonic techniques furthermore make it possible to estimate marginal prices for individual characteristics of the bundled product. For an example of this technique, see Cole, et al (1986).

3 What is Prepackaged Software?

Prepackaged Software is Industry 7372 in the 1987 United States Standard Industrial Classification (SIC). "Prepackaged" refers to software typically bought "off the shelf," in other words, not custom software. The industry includes establishments primarily engaged in software design, development, and production. Of these activities, design and development command the lion's share of input costs. The product of the industry is actually the lines of code that make up a computer program. A licence, not the program itself, is typically sold which: a) gives the purchaser the "right to use" the program, and b) gives the company leverage to combat software piracy. The program itself cannot be altered or customized. Custom software is in SIC 7371, Computer Programming Services.

The international System of National Accounts (SNA, 1993) treats computer software as intangible fixed assets, or investment goods.

3.1 Varieties of Prepackaged Software

Prepackaged software products can be divided into: a) operating and utility programs, and b) applications programs. Operating and utility programs, known as "systems" software, are used to control, manage, monitor, or enhance a computer's capabilities. Applications programs are executed by the computer to perform specific functions, such as word processing, spreadsheet data manipulation, data base operations, financial analysis, and so on.

3.2 How big is Prepackaged Software, Applications Software in particular?

In 1995, the prepackaged software industry employed 181,000 individuals, a 16% increase over 1994, and more than double the level of 1988 (Freeman, 1996). Between 1993 and 1994, sales of applications programs for entertainment, education, and graphics grew faster than sales of word processor, spreadsheet, and data base programs, roughly mirroring growth differences between personal computers purchased by families versus those purchased by businesses. In dollar terms, Census estimates[1] that SIC 7372 receipts for 1992 were $20.8 billion, and Oliner & Sichel (1994) cite that Software Publishers Association (SPA) sales estimates for spreadsheet, word processor, and data base applications programs totaled $5.7 billion in 1992 and $6.8 billion in 1993. The average annual growth of these three types of applications programs was 20.1% from 1987 to 1993.

---

[1]U.S. Department of Commerce. 1992 Census of Service Industries. Subject series: Computer and Data Processing Services.

4 Literature on Hedonic Price Analysis

4.1 Beginnings

The literature on hedonics is voluminous. It's beginnings can be tracked to studies by Waugh (1928 and 1929) in which he systematically graded vegetables at market according to certain perceived desirable quality characteristics, and then recorded bundled lot prices. Prices were compared to characteristics and developed into a fairly crude hedonic model used to describe what premiums buyers were willing to pay for certain characteristics. It was assumed that rational, profit-maximizing farmers would use this information and do their best to satisfy customer wants.

Court (1939) applied the concept of hedonic regression analysis to car prices to demonstrate that although nominal prices for automobiles indeed increased rapidly between 1925 and 1935 (45%), when adjusted for quality improvements, they actually dropped over that period. At the time, the United States was pondering whether it should try to stabilize production and employment in this huge industry by manipulating car prices. Court described hedonic price comparisons as "those which recognize the potential contribution of any commodity, a motor car in this instance, to the welfare and happiness of its purchasers and the community."

After a long hiatus in this branch of econometric work, hedonic price analysis was revived by Griliches (1961) and others. In addition to price applications, hedonics theory is often seen today in litigation cases involving loss of life, loss of income, etc.

4.2 Prior Hedonic Research in Prepackaged Software

Identifying prior hedonic research in the area of software products yielded two studies, Brynjolfsson & Kemerer (1993) and Gandal (1994), and these dealing only with spreadsheet software. Brynjolfsson's data are not available (as it is being used in a software litigation case), but their primary source for prices and characteristics, National Software Testing Labs (NSTL) *Software Digest Ratings Reports*, was obtained. In addition, Oliner & Sichel (1994) used traditional matched-model techniques to estimate average annual price changes for spreadsheet, word processing, and data base programs. They estimated an average annual price decline of 4.0% over the period 1987-1993. The Brynjolfsson & Kemerer study estimated average an annual average price decline of 6.4% over the period 1987-1992. Gandal's model showed an average annual price decline of 15% for the period 1986-1991.

The large difference in the results of the hedonic studies so far remains unexplained, beyond differences in source data and functional form. Gandal's primary source for data is DataPro, and he uses a log-linear function, whereas Brynjolfsson & Kemerer's primary source is *Ratings Reports*, and uses a linear function[2].

5 The Hedonic Pricing Model

Basically, the hedonic model regresses product prices on product features or "characteristics" over time. The goal is to develop an index which when factored into the nominal series, yields real prices, holding quality constant. The crucial assumption to the theory is that an item and it's price represent a bundle of many features, each of which contributes in some way to the value of the item or represents a cost to its development. Diverse characteristic features cumulatively define the product, and choosing which are "representative" is a subjective endeavor, limited in part to what data is available.

The model in notation:

(1) $$P = f(X)$$

(2) $$f(X) = \alpha_0 + \beta_j FEATURE_j + \beta_t YEAR_t + E$$

Where

P = list prices, or the natural log of list prices

β_i = regression coefficients for a list of j product features

β_t = regression coefficients for a list of t years[3]

E = error term

There is no agreed "best" functional form. Choices include linear, log-linear, and log-log. If linear, the list price and product characteristics are regressed as is. If log-linear, either the price or characteristics are first transformed into natural log values before entering the regression. If log-log, both sides of the equation are transformed into natural logs pre-regression. Whichever form yields the best goodness of fit is usually chosen.

---

[2]Functional form is discussed in the Hedonic Pricing Model section following.

[3]Wallace (1996, p. 22) interprets this as "The regression parameter measuring period *t* residual mean price change between periods once the mean changes in characteristics costs have been accounted for."

Gujarati (1988) interprets the slope coefficient of the log-linear function (when the dependent variable alone is transformed into natural logs) as a measure of:

> "The constant proportional or relative change in Y for a given absolute change in X...This functional form is particularly useful in situations where the X variable is a time trend, because then the model describes the constant relative or constant percentage rate of growth or decay in the variable Y."

The partial regression coefficients of such log-linear functions represent the percentage increase/decrease for a unit change in the characteristic, taking into account the effect of other independent variables. For example, suppose a variable takes on the value of 0 or 1, and it's regression coefficient is +0.10. The bundled product can expect to command a price premium of 10% over similar products in which this particular feature is absent.

In a log-log specification, partial regression coefficients can be interpreted as elasticities of price with respect to product features.

Whatever function the equation takes, when separate time dummies are used (excluding the base year), an index can be built directly. Berndt (1991, pp. 122-123) offers a direct and simple procedure involving taking the antilogarithms of the estimated coefficients of the dummy variables. For instance, if the first year is standardized to one, and the second through fourth years time dummy coefficients are -0.1398, -0.4891, and -0.5938, the actual index is:

Year 1	1.0000
Year 2	0.8695
Year 3	0.6132
Year 4	0.5522

This procedure is used in this study.

6 Procedures for this study

Data bases were built for both spreadsheet and word processing software, and are used as the basis for all regressions. Described first will be results of the spreadsheet regressions, followed by the word processor regressions. As noted at the beginning, the spreadsheet methodology attempts to follow that of previous studies, the purpose being to verify or confirm that body of work. Word processor regression methodology took a new approach, in that variables pre-selected as candidates to represent desirable characteristics were developed methodically based on numerical rankings found in late issues of *Ratings Reports*. Study results are based exclusively on NSTL *Ratings Reports* data, and cover the period 1985-1994. DOS, Window, Mac, and OS/2 operating platforms are included.

One would expect NSTL's *Ratings Reports* to have a distinct advantage over DataPro (Gandal's source), in that while both *Ratings Reports* and DataPro follow a fairly standard format year-to-year, thereby facilitating comparisons over time, *Ratings Reports* tracks a considerably broader range of features.

6.1 Spreadsheet Software

6.1.1 Spreadsheet Software Data Considerations

The explanatory variables chosen are the same as those selected by either or both Gandal and Brynjolfsson, and for which fairly complete data was available back to 1986[4].

Brynjolfsson & Kemerer chose descriptive variables based "upon the importance placed upon them in contemporary reviews." Which reviews are not specified. Gandal chose descriptive variables that were "available" from DataPro, indicating a more objective search criteria.

I make the assumption that when going back in time, if a variable ceases to be discussed, the technology is considered non-existent, which is consistent with the treatment used to develop the word processing software portion of this study. Doing so

[4]Both previous studies used vintage dummy variables in order to run specification tests per Berndt and Griliches (1993). If a model is correctly specified, vintage dummy variables are expected to be insignificant. Beginning in 1990 however, *Ratings Reports* fail to track introduction date consistently. As this study relies exclusively on *Ratings Reports* data, it is left for further study to develop such variables. Both previous studies however reported that including vintage dummies had an insignificant effect.

assigns values of 0 (not present) and yet inclusion in the model, rather than 9999 (not available) and exclusion.

Both linear and log-linear functions were tested, with summary statistics in each case very close. For expository reasons I chose a log-linear functional form, in that the list prices are translated first into their natural logs while characteristics values, with the exception of MIN_RCLG[5], are not.

Observations are unweighted. Gandal used unweighted observations, and Brynjolfsson included an "Installed base" variable, which is the sum of its sales in all prior years, including sales of earlier, compatible versions, divided by the sum of the installed base for all products. His sources for estimating the installed base were DataQuest and International Data Corporation. Explanatory variables include:

CELLLINK	A cell reference from an external worksheet can be used in a formula in the current worksheet.
COL_SORT	The option to sort by columns.
EMBED_CH	Charts may be embedded directly in the worksheet.
EXT_LINK	Linking capability with dBase files.
GRAPHS	Can produce pie, bar, and line graphs.
KEY_MACRO	Macros can be written in a “learn mode”, where the spreadsheet can convert keystroke steps into a program.
LOT_COMP	Compatible with .WKS or .WK1 formats. Capable of exchanging files with Lotus spreadsheets.
LOT_PROD	Product of Lotus Corporation.
LOT_TREE	An exact duplicate of the Lotus 1-2-3 menu tree, or a menu tree that operates in the same manner.
MACROS	If-Then-Else programs can be used to automate repetitive tasks.
MIN_RCLG	The minimum of the maximum number of rows and columns that the spreadsheet can handle, in natural log form.
PRESFEAT	Worksheets and graphs can be printed on the same page, and/or multiple printing fonts and character sizes are available. Value = of 0, 1, or 2.
PRINFEAT	Takes on the value of one if three or more of the following five print functions are possible: Sideways printing, background printing, preview mode, PostScript support, and printing of noncontiguous worksheet portions.
SORT>=2	Can sort a group of data observations on at least two levels.
WINDUM	Maximum number of windows on screen simultaneously. Dummy=0 if less than two, 1if between two and fifteen, inclusive, and 2 if greater than 15.

[5]Gandal first translated MIN_RC into log form, because actual values ranged from 63 to 32,800. Both forms were tested for this study. While neither form yielded a statistically significant coefficient, using the log form yielded a non-zero coefficient, and hence is used in regressions two and three.

WYSIWYG	Offers a What-You-See-Is-What-You-Get interface.
TIMEDUMMIES	One for each year, excluding the first year 1986. Even though a 1994 issue of Ratings Reports is available, and another index year would be desirable, this study chose to ignore it."[6]

Regressions are run primarily to compare results with previous studies. For instance, Regression 1 includes only variables used by Brynjolfsson & Kemerer, and Equation 2 includes only variables used by Gandal. Regression 3 uses variables used in both studies, and Regression 4 includes variables used in both studies, but only if data is available for all observations[7].

6.1.2 Spreadsheet Software Regression Results

For the four regressions, estimated average annual compound price change ranged from -10.9% to -16.9% from 1986-1992, which place them somewhere closer to Gandal. When looking at the full period 1986-1993, the price change ranges from -9.0 to -13.4. The largest price declines occur in the early years, as new technology features emerge (EMBED_CH and WYSIWYG to name two) in products even as list prices remain relatively stable. Common across regressions is an apparent premium assigned to Lotus Corporation products, or products that have a menu tree that either looks quite like a standard Lotus 1-2-3 menu tree, or operates in the same manner. In Regression 1 for instance, the coefficient for LOT_TREE is interpreted to mean the package would typically command a 52% higher price than those without, holding other variables constant. Significant premiums are placed likewise on products which allow greater numbers of windows to be displayed, or feature macro "learn mode" capability.

While non-time variable coefficients are intuitively expected to be positive, this is

---

[6]According to NSTL, issue (11.8, 1994), Dynamic Viewing Spreadsheets, "Differ in scope in offering dynamic numeric data viewing capability...providing users flexible viewing of data by simply rearranging a preset category or group. This grows out of the need to create reports using the massive amount of data accumulated by, for instance, multinational corporations." Although Dynamic Viewing allows analysis, modeling, and reporting, many features normally discussed are absent. NSTL suggests that "for a full discussion on the many features available in spreadsheets please consult Software Digest 10.12 (1993), Advanced Spreadsheets." Issue (10.12) is included in this study.

[7]This regression may be viewed as 'observation-inclusive,' at the expense of explanatory variables COL_SORT, EXT_LINK, MIN_RCLG, PRINFEAT, WINDUM, and WYSIWYG, which each contain not-available observations for some issues.

not always the case. In all but one instance[8] however, these have insignificant t-values.

Table 1 and Figure 1 show results of the four spreadsheet software regressions. R-Square is .463 for Equation 1 (emulating Brynjolfsson & Kemerer's .553) and .628 for Equation 2 (emulating Gandal's .857). The rather large difference in explanatory power between this study and Gandal's can likely be attributed both to the use of different source data, as mentioned, and to the necessary occasional judgmental interpretations of quality characteristics from *Reports* issues.

Looking at these independently derived results, one can be fairly confidant that previous hedonic studies, which show significant price decreases over time while allowing for quality improvements, are correct. Less obvious is how dramatic that price decline is.

6.2 Word Processing Software

One valid criticism of hedonic work, noted earlier, is the subjective nature of how descriptive variables are chosen. This segment of the study takes an objective, and different, approach in that it proceeds methodically to determine included variables from a rather long list of candidates.

The most recent (February 1994) issue of *Ratings Reports* is used as the 'defining' issue, from which a representative set of characteristics are chosen. Overall, 341 product features are tracked in this issue by NSTL, an unwieldy number for a study such as this. To narrow the set of characteristics to be included, each of 26 broad general categories (Editing Functions, Paragraph Formatting, etc.) was reviewed. Within each category, features having the highest weight, as determined by NSTL[9], are selected. When several features tie for the highest weight in a category, each of these is included. If all features in a category have the same weight, a previous *Report* is referred to for a tie-breaker because earlier reviews typically had unequal weights for the features. Sometimes it is necessary to go back several issues to find a tie-breaker. Table 3 lists and describes the variables chosen using the above rules.[10]

---

[8]The variable SORT>=2. Gandal also found this variable to have a negative coefficient, and simply excluded from most regressions.

[9]The Editing Functions category, for instance, tracks 18 product features, whereas this study tracks only those ranked highest; in this case a three-way tie among ABBREVGL, DRAGDROP, and NUMBUNDO.

[10]When necessary to go to previous issues to find a tie-breaker, the *Report* issue number is noted in the table.

6.2.1 Word Processing Software Data Considerations

Certain problem data considerations are worth mentioning. Some variables involve considerable judgement the further away they are reviewed from the initial 1994 issue. For example, TOLLFREE and FREETECH are specifically rated in the 1994 issue, but must by necessity be determined in earlier issues by reviewing each package's summary page, a tedious exercise yielding not always clear results. If an issue specifically cites one package that includes a toll-free number, but is moot on other packages, I necessarily assume the others do not. If technical support is mentioned in some and not mentioned in others, I must judge whether that support is free or for a fee. Sometimes, free technical support is available for a rather short period (30 days, for example). Does one consider that free?[11]

For the Printing major category, one must go back to August 1991 (8.9) to find a tie-breaker to decide which of several variables to include. Doing so yields two, BACKPRIN and PRINQUEU. But the December 1992 issue evaluates neither, and the February 1994 issue mentions only BACKPRIN. This is a dilemma. For word processing software, "Printing" as a subgroup is valued highly by NSTL raters (a weight of 4, the highest, in issues leading up to February 1994, when the subgroup is downgraded to a still-respectable weight of 2). This study assumes that for the December 1992 issue, BACKPRIN and PRINQUEU are not available (=9999) in all cases.

Given the problems mentioned, ASCII, TOLLFREE, FREETECH, BACKPRIN, and PRINQUEU, are excluded from the expanded-variable regression. BACKPRIN and PRINQUEU are excluded because by including them the four observations of lower-priced packages from rating year 1992[12] all must drop out, a critical loss.

While a major object of this part of the study is objective pre-selection of variables based on highest weight, one can see that the selection of a suitable variable set is not always straight-forward. An effort is made to include as many 'highest weight' features as possible.

---

[11]Take the variable ASCII as another problem variable. While specifically rated in the 1994 issue, it too is only found in earlier issues from scanning each package's summary page. Often only ASCII "write" ability is specifically mentioned. Does one assume "read" ability is also present? Doing so yields a vector of 1's-- universally present in all packages in all years--and the variable ASCII becomes meaningless.

[12]The December 1992 issue of *Ratings Reports* features "Executive," or lower-priced programs.

6.2.2 Word Processing Software Regression Results

The data set for word processor software has 155 observations. The dependent price variable is transformed into natural logs, based both on testing (summary statistics are slightly better when using the transformed price) and for comparison purposes with spreadsheet software results. Regression 1 is 'observation-inclusive' in that it includes twenty non-time variables which have data available for all *Reports*. Regression 2 is 'variable-inclusive' in that 25 additional variables are included, at the expense of observations lost for lack of available data in certain *Reports*.

As expected, estimated prices declined over the period of study. Average annual price change in Regression 1 (the 20-variable case) is -15.1% , with prices declining steadily over the 1985-1994 period. For Regression 2, the entire price decline occurs between 1985 and 1991, with prices remaining quite stable 1991 forward. Over the period 1985-1994, the annual average price change is -18.5%.

Apart from the downward march in product prices over time, inspection of the qualitative variables reveals some interesting results. Of variables included in both regressions, MAILMERG and HYPHSOFT appear to be most valued. For instance, the ability to merge a data file containing names or addresses into a document to produce a series of personalized form letters commands a 70% premium in price over other word processing packages, holding all other variables constant. A program's ability to automatically remove a hyphen after editing places the entire word on one line receives a similar price premium of approximately 30%. Furthermore, programs featuring the ability to edit a macro or to edit a program-supplied dictionary are also highly valued, commanding approximate price premiums of 40% over programs lacking these features.

One would expect the sign of the partial coefficient for all characteristics variables to be positive, but this is not the case. Five of twenty (20%) non-time variables in Regression 1 have negative coefficients, but only one of these (AUTOSAVE) is statistically significant. This percentage jumps to 47% (21 of 45 variables) in Regression 2, with three formerly positively-related coefficients turning negative. In Regression 2, no negative coefficients are statistically significant however. Wallace (1996) makes note of one of the drawbacks for estimating hedonic-based price functions, that of selecting the "correct" set of characteristics. Are these the "correct" set of characteristics? While including all available highly-rated features may be theoretically appealing, results point to further work in selecting a representative variable list. Regression results for Word Processing Software are shown in Table 2 and Figure 2.

7 Test for Structural Change

In studies such as this, which attempt to track the direction and level of prices for high-technology goods, it is especially important to address the following question: Does the model reliably represent the entire period of study? Or is there a basic shift over time, making early/late periods not statistically comparable? Do the characteristics for one time period adequately represent the typical bundled product in the face of rapid technological improvements?

Regression results reported in this study are derived by pooling all years together. Doing so makes the assumption, and imposes the restriction, that the slopes of the subperiods are equal to each other and to the pooled time period. A CHOW test is run to support or refute the hypothesis: The relationship among the variables in the model remains constant over time. If the model passes this test for robustness, one is more confident the regression results hold true not only for the period tested, but for future periods as new data become available.

7.1 Test for Structural Change in the Spreadsheet Software Data Set

For prepackaged spreadsheet software, the test is run using the following two criteria: (1) use of "best-available" variables, i.e. those for which data in subperiods 1986-1988 (41 observations) and 1989-1993 (45 observations) did not contain vectors of all 1's or 0's,[13] and (2) use of "observation-inclusive" variables.[14] The CHOW F-statistic of 2.1493 (2.0741 table) barely rejects the hypothesis at the 5% confidence level, and cannot reject the hypothesis at the 1% confidence level (2.7771 table).

One cannot say with a high degree of certainty therefore that the two time periods are much different. Indeed, using a similar but different set of years, one cannot reject the hypothesis at all[15]

---

[13]This was a particular problem in the early period, as technology for several variables simply did not exist. Variables are Celllink, Key_macr, Lot_prod, Lot_tree, Macros, Presfeat, Ratingyr, and the constant.

[14]Those that have no not-available observations over the pooled period, in order to maximize the number of observations.

[15]A CHOW test on a second split, 1986-1989 (52 observations) and 1990-1993 (34 observations) failed to reject the hypothesis at both 5% and 1% confidence levels: 1.2711 (2.0210 table 5%, and 2.6801 table 1%)

7.2 Test for Structural Change in the Word Processing Software Data Set

For prepackaged word processing software, two tests are run, again using the "best-available" and "observation-inclusive"[16] variable criteria. The first test splits roughly on equal number of observations (1985-1987; 80 observations vs. 1988-1994, 75 observations), and the second on roughly equal number of years (1985-1988, four years vs. 1989-1994, six years)[17]. In both cases, the computed statistic could not reject the hypothesis that the relationship among the variables in the model remain constant over time. For the first, the test statistic is 1.313 (1.7261 table 5%, and 2.1511 table 1%). For the second, the test statistic is 0.9656 (1.6631 table 5%, and 2.0432 table 1%).

[16]After satisfying "best-available" criteria, remaining word processing software variables have no not-available values, so satisfying this criteria does not lead to a loss of observations.

[17]Splitting on equal number of years would result in an unworkable regression, in that the 1985-1989 period has 135 observations, and the 1990-1994 period only 20.

8 Conclusions

- It is concluded that prior hedonic price studies for prepackaged spreadsheet software correctly show a decline in prices during the late 1980's and early 1990's. In addition, one may look to this verifying study for information on the relative value users place on specific product features.

- It is further concluded that by using objective criteria for pre-selecting product characteristics, a reliable econometric model can be built to determine hedonically the movement in prepackaged word processing software prices over the same time period.

References

Berndt, Ernst R. The Practice of Econometrics: Classic and Contemporary. MIT and the National Bureau of Economic Research. Addison-Wesley Publishing Company, 1991.

Berndt, Ernst R. and Zvi Griliches. "Price Indexes for Microcomputers: An Exploratory Study." In Murray F. Foss, Manser, and Young, eds. Price Measurements and Their Uses, Studies in Income and Wealth. Conference on Research in Income and Wealth, Volume 57. Chicago: University of Chicago Press for the National Bureau of Economic Research. 1993: 63-93.

Brynjolfsson, E. and Chris F. Kemerer. "Network Externalities in Microcomputer Software: An Econometric Analysis of the Spreadsheet Market." Center for Information Systems Research, Sloan School of Management, Massachusetts Institute of Technology. CISR Working Paper No. 265. November 1993.

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

Commission of the European Communities, International Monetary Fund, Organization for Economic Co-operation and Development, United Nations, and World Bank, System of National Accounts 1993 (Brussels, Luxembourg, New York, Paris, Washington, D.C.: Commission of the European Communities, International Monetary Fund, Organization for Economic Co-operation and Development, United Nations, and World Bank, 1993).

Court, Andrew T. "Hedonic Price Indexes with Automotive Examples." In The Dynamics of Automobile Demand, New York: The General Motors Corporation. 1939: 99-117.

Freeman, Laura. "Job Creation and the Emerging Home Computer Market." Monthly Labor Review. August 1996: 47-56.

Gandal, Neil. "Hedonic Price Indexes for Spreadsheets and an Empirical Test for Network Externalities." RAND Journal of Economics, Vol. 25, No. 1. Spring 1994: 160-170.

Griliches, Zvi. "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, 1961: 137-196.

Oliner, Stephen D., and Daniel E. Sichel. "Computers and Output Growth Revisited: How Big Is the Puzzle?" Brookings Papers on Economic Activity, 2: 1994: 273-334.

Triplett, Jack E. "The Economic Interpretation of Hedonic Models." SURVEY of Current Business, 66:1, January 1986: 36-40.

Wallace, Nancy E. "Hedonic-Based Price Indexes for Housing: Theory, Estimation, and Index Construction." Economic Review, Federal Reserve Bank of San Francisco. 1996: 34-48.

Waugh, Frederick V. "Quality Factors Influencing Vegetable Prices," Journal of Farm Economics, 10:2, April 1928: 185-196.

_____. "Quality as a Determinant of Vegetable Prices: A Statistical Study of Quality Factors Influencing Vegetable Prices in the Boston Wholesale Market." New York: Columbia University Press. 1929. Reprinted by the AMS Press, Inc. New York, 1968.

Table 1-1: Regressions of Prepackaged Spreadsheet Software

Variables	Coef.	Coef.	Coef.	Coef.
CELLLINK	0.366 *	0.024	-0.025	0.288
COL_SORT	0.385 *		0.301	
EMBED_CH	0.500		-0.065	0.028
EXT_LINK		-0.214	-0.251	
GRAPHS		0.363	0.211	0.340
KEY_MACR		0.415 *	0.406 *	0.565 **
LOT_COMP		-0.043	-0.079	0.045
LOT_PROD	0.463 *	0.630 **	0.796 **	0.395 *
LOT_TREE	0.524 **		0.330 *	0.436 *
MACROS		-0.198	-0.069	-0.021
MIN_RCLG		0.033	0.028	
PRESFEAT		0.271 *	0.245	0.331 *
PRINFEAT		0.384 *	0.296	
SORT>=2		-0.457	-0.587 *	-0.571 *
WINDUM		0.358 **	0.362 **	
WYSIWYG	0.434		0.207	
YEAR87	-0.191	-0.103	-0.188	-0.315
YEAR88	-0.274	-0.142	-0.232	-0.249
YEAR89	-0.967 *	-0.440	-0.719	-0.676 *
YEAR90	-0.858 **	-0.673 *	-0.724 **	-0.718 **
YEAR91	-0.810 *	-0.423	-0.565	-0.563
YEAR92	-1.110 **	-0.690	-0.841 *	-0.762 *
YEAR93	-1.004 *	-0.658	-0.779	-0.703
Constant	5.102	4.933	4.875	4.945
Equations	(1)	(2)	(3)	(4)
MultipleR	0.680	0.792	0.828	0.744
RSquare	0.463	0.628	0.685	0.554
AdjustedRSQ	0.352	0.508	0.549	0.443
StandardError	0.554	0.477	0.462	0.492
ResDF	63	59	53	68
ResSUMSQ	19.343	13.441	11.333	16.435
ResMeanSQE	0.307	0.228	0.214	0.242
F	4.177	5.246	5.018	4.974
Signif. F	0.0001	0.0000	0.0000	0.0000
Cases	77	79	77	86
Durbin-Watson	2.034	2.094	1.887	2.042
Chg: 1986-1992	-16.9	-10.9	-13.1	-11.9
Chg: 1986-1993	-13.4	-9.0	-10.5	-9.6

(*) Coefficient is significant at 5%. (**) Coefficient is significant at 1%.

Table 1-2: Regressions of Prepackaged Word Processing Software

Variables	(1) Coef.	(2) Coef.	Variables	(1) Coef.	(2) Coef.
ABBREVGL		-0.314	MAILMERG	0.684 **	0.710 **
ALT_HEAD	0.122	-0.058	MARKINSE		-0.002
AUTOSAVE	-0.257 *	-0.219	MAXIOUTL	-0.000	0.001
AUTOBACK		-0.036	MULTCHAR	0.011	0.092
AUTOEQUA	0.273	0.150	NUMBUNDO	0.003	0.001
AUTONUMB		-0.018	PAGENUMB		-0.184
CASESENS		-0.251	PASSWORD	0.323	0.239
COLPAGLN		0.181	READONLY		-0.116
COLUFORM		0.373	SEARBLOK		0.108
CONDMAIL		-0.107	SELEMAIL	-0.210	-0.010
DRAGDROP		1.893 *	SHOWHIDE	0.333	0.382
EDITDICT		0.420 **	SINGCHAR	-0.082	-0.100
EDITMACR	0.440 *	0.229	STEPDEBU		-0.376
FOOTENDN		-0.071	STYLFILL	0.126	-0.872
FREETECH	0.085		TABLCONT		0.322 *
GENECHAR		0.219	TEXTGRAF		0.801
GRAMCHEK	0.108	-0.129	TOLLFREE	0.245	
HEADFOOT		0.274	USERSPEC		-0.474
HYPHAUTO	0.585 **	0.345	VARIWIDT		0.176
HYPHHARD		-0.176	YEAR86	-0.157	-0.294 *
HYPHSOFT	0.274 *	0.334 **	YEAR87	-0.271	-0.501 *
ICONPALT		0.026	YEAR88	-0.341	-0.853 **
IMPSPRTB		-0.761	YEAR89	-0.641 *	-0.989 **
KEEPTEXT	-0.125	-0.228	YEAR91	-0.702 *	-2.252 *
LEARNMOD	-0.067	-0.140	YEAR92	-1.232 **	-2.305
LOCKFILE		0.408	YEAR94	-1.475	-1.841
			Constant	4.795	4.881

Equations	(1)	(2)		(1)	(2)
MultipleR	0.775	0.878	ResDF	127	90
RSquare	0.600	0.770	ResSUMSQ	32.59	18.19
AdjustedRSQ	0.515	0.643	ResMeanSQE	0.257	0.202
StandardError	0.507	0.450	F	7.065	6.032
RegDF	27	50	Signif. F	0.0000	0.0000
RegSUMSQ	48.96	60.97	Cases	155	141
RegMeanSQE	1.81	1.22	Durbin-Watson	1.776	1.936
Chg:1985-1994	-15.1	-18.5	Signif: (*)=5%, (**)=1%		

Table 3: Word Processor Program Variables

CHARACTERISTICS based on **categories** in the February 1994 issue (11.2) of *Ratings Reports*

Editing Function
DRAGDROP. Users can mark text and move or copy it to another location by moving the mouse pointer.
NUMBUNDO. The number of most recent commands stored by the program to allow multiple-step undoing.
ABBREVGL. A program function allows entry of selected boilerplate text with the use of an abbreviation or command.

Editor Display characteristics
COLPAGLN. The position of the current cursor location displays on-screen, giving page, column, and line numbers.
HEADFOOT. Editing mode allows editing and display of leaders, footers, and footnotes as they will appear in the final printout.
USERSPEC. Users can specify a percentage of actual size for viewing a document. Larger than 100% allows magnification; smaller allows more of the document to be seen at a time.

Search and Replace Functions
CASESENS. The program can optionally be sensitive to, or disregard, the case of characters when performing a search. For example, specifying "otter" will find "otter," "Otter," and "OTTER."
SEARBLOK. The program can search a user-specified text block instead of the entire document.
SINGCHAR. The program can find variations of words based on a wildcard search character for which the program substitutes a single character; for example, specifying "t?n" will find "ten," "ton," and "tin".
MULTCHAR. The program can find variations of words based on a wildcard search character for which the program substitutes multiple characters; for example, specifying ; "t*n" may find "ten," "then," and "tendon."

Revision Marking
MARKINSE. Text can be marked with a special character that when printed (e.g., a line through it) shows that it should be deleted or inserted. The program can automatically remove or insert text that has been so marked.

General Formatting (from 9.10, Dec. 1992).
A section of text (e.g., a paragraph or table) can be kept together on a single page regardless of the natural page break.

Paragraph Formatting
BULLETNG. A paragraph formatting option causes each new paragraph to begin with a bullet.
AUTONUMB. A program option causes each new paragraph to be numbered one greater than the previous.
AUTORENU. When a numbered paragraph is moved or deleted, the program automatically renumbers the affected paragraphs.

Style Sheets/Templates/Forms
USESTYLE. Once styles have been defined, they can be saved for use in any document.
NUMBTEMP. The number of templates and forms shipped with the program.

Hyphenation (all from (9.10) December 1992).
HYPHSOFT. The program can automatically remove a hyphen after editing or reformatting causes the word to move such that it is no longer divided over two lines.

HYPHAUTO. The program either hyphenates or suggests hyphenation for words that extend beyond the right margin.
HYPHHARD. A word that should always be hyphenated (e.g., a compound word) can be hyphenated with a hyphen that remains regardless of whether the word is broken at the end of a line.

Column Formatting
VARIWIDT. Users can specify newspaper columns of varying widths and of varying amounts of space between them.
COLUFORM. The program can mix multicolumn formats on the same page (e.g. centering one heading over three columns of text).

Tables and Math
STYLFILL. Users can choose line styles and fill patterns for tables.
TEXTGRAF. Entries in cells can include graphic objects as well as text.
IMPSPRTB. The program can import data directly from a spreadsheet program into a table.

Headers and Footers (both from (9.10) December 1992)
ALT_HEAD. Users can specify that one header be printed only on even pages and that a different header be printed only on odd pages of a document (e.g., book chapter heading on the left page and subsection heading on the right).
PAGENUMB. Users can select the page number that will be assigned to the first printed page.

Footnotes and Captions
FOOTENDN. The program can place footnotes at the bottom of a page, or collect them as endnotes at the end of the document.

Index Contents and References (from (9.10) December 1992)
TABLCONT. The program can create a table of contents from user-marked references or from heading styles.

Spelling/Thesaurus
EDITDICT. Users can delete words from, and add words to, the user dictionary.

Mail Merge (from (9.10) December 1992)
MAILMERG. The program can merge a data file (e.g., containing names and addresses) into a document to produce a series of personalized form letters.
CONDMAIL. Users can specify that selection criteria be applied during a merge. For example, depending on the account due date printed in a document, the program may insert a paragraph explaining that the account is overdue and request immediate payment.
SELEMAIL. Mail merge options include the ability to select only a subset of the data file (e.g., only those customers whose last names begin with the letter "S").

Outlining (from (9.10) December 1992)
MAXIOUTL. The maximum number of levels that can be included in an outline.
SHOWHIDE. All subheadings below a user-specified level can be expanded or hidden.

Drawing and Graphics Support
DRAWPLAC. Users can draw objects and import pictures directly on a document with no need to draw or import in a separate screen.
CLIP_ART. The program ships with clip art objects.

Charting/Scientific Equations (from (9.6) September 1992, pp.51-52)
GENECHAR. The program can generate pie or bar charts using data from a table in the document or in a spreadsheet.
AUTOEQUA. The program automates the process of creating symbols that operate on multiline equations, which can be entered in the middle of a line of text.

Printing (from (8.9) August 1991, p.60)
BACKPRIN. The program can print documents while the user works in the foreground.
PRINQUEU. The user can specify that a series of different documents be printed one after another.

File Handling (from (9.10) December 1992)
AUTOBACK. When users begin to re-edit a document, the program automatically creates a copy of the original file and saves it to disk.
AUTOSAVE. The program automatically saves the current file during an editing session, or at intervals set by the user.

Document Summary Information
No weighting found greater than one for any issues going back to (5.7) July 1988, when this characteristic group does not appear at all. Dropped.

Programming/Macros
EDITMACR. A macro definition can be edited as a document.
LEARNMOD. The program can automatically record a user's keystrokes during a work session and assign them to a macro name or key.
STEPDEBU. The macro function includes a debugging option of macro playback one step at a time (aids in finding the source of problems with macros).

Customizing (from (9.10) December 1992)
ICONPALT. Users can modify program icons (adding or deleting items) to fit specific needs.

Data Import/Export
ASCII. The program can read and write files in ASCII and ASCII delimited format.

E-Mail/Networking
LOCKFILE. A file is locked when one user edits it; no other user can edit the same file simultaneously.
READONLY. The program's editor can load and display files to which users have read-only privileges.

Miscellaneous Features
GRAMCHEK. The program includes a grammar checker that can be used to check common misuse of noun-verb agreement, tense, cliches, and others.
PASSWORD. The program provides password protection of documents.
FREETECH. The vendor does not charge a fee to registered users for technical support.
TOLLFREE. The vendor offers a toll-free telephone number to registered users for technical support.

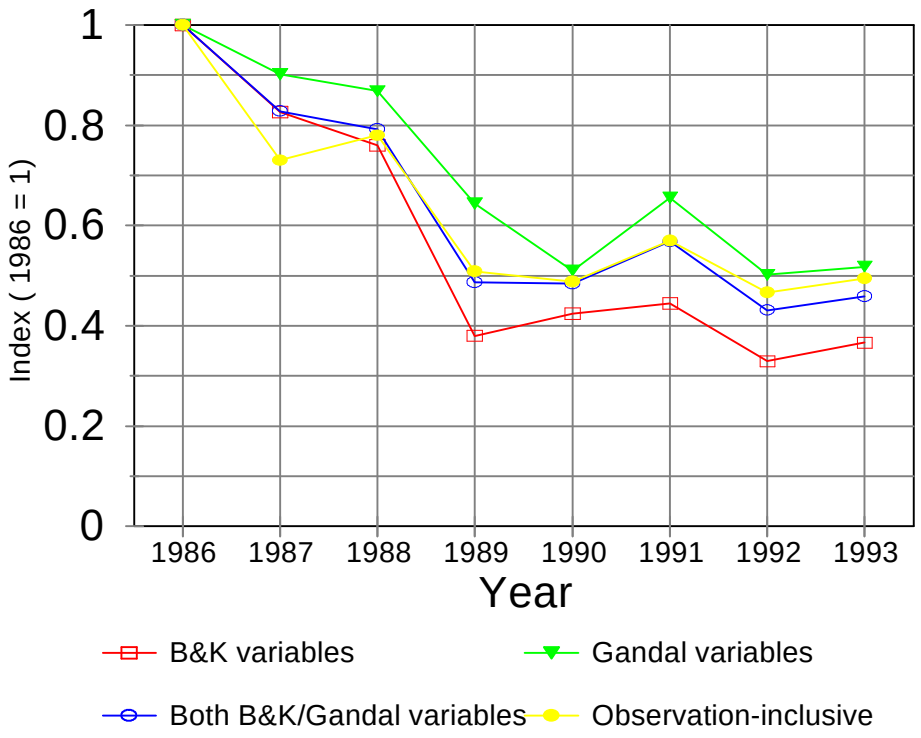

Figure 1: Spreadsheet Software

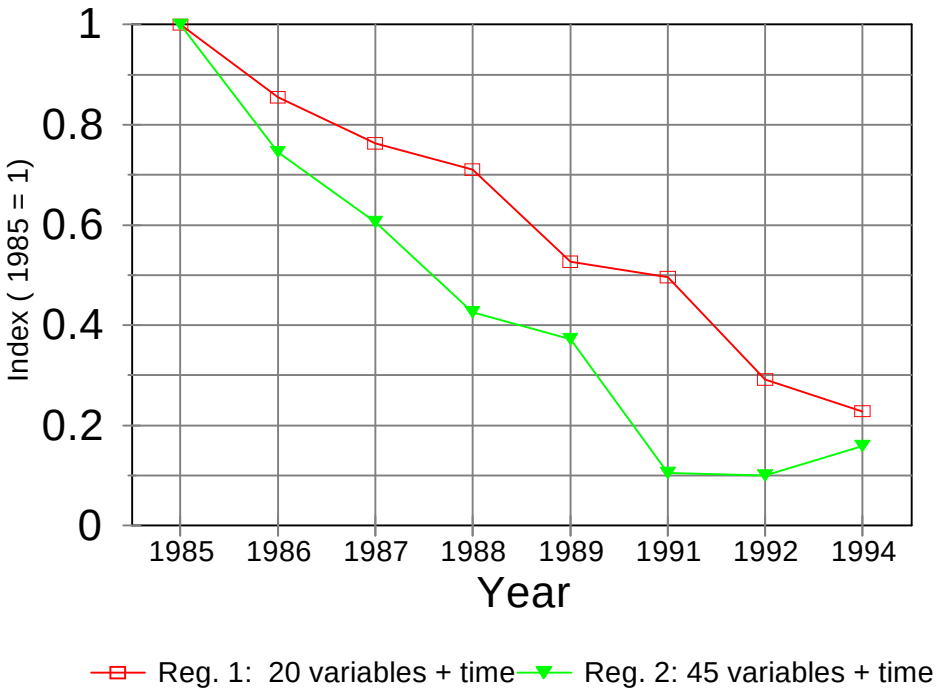

Figure 2: Word Processing Software

Vita

The author was born October 26, 1952 in Washington, DC. After graduating from George Washington University with a bachelors degree in Business Administration concentrating in economics and public finance, he began and continues a career as an economist with the U.S. Department of Commerce.