Network Externalities in Microcomputer Software: An Econometric Analysis of the Spreadsheet Market

Erik Brynjolfsson • Chris F. Kemerer
MIT Sloan School of Management, E53-313, 50 Memorial Drive, Cambridge, Massachusetts 02142
Katz Graduate School of Business, University of Pittsburgh, 278A Mervis Hall, Pittsburgh, Pennsylvania 15260

Because of network externalities, the success of a software product may depend in part on stalled base and its conformance to industry standards. This research builds a hedonic model to determine the effects of network externalities, standards, intrinsic features and a time trend on microcomputer spreadsheet software prices. When data for a sample of products during the 1987–1992 time period were analyzed using this model, four main results emerged: 1) Network externalities, as measured by the size of a product's installed base, significantly increased the price of spreadsheet products: a one percent increase in a product's installed base was associated with a 0.75% increase in its price. 2) Products which adhered to the dominant standard, the Lotus menu tree interface, commanded prices which were higher by an average of 46%. 3) Although nominal prices increased slightly during this time period, quality-adjusted prices declined by an average of 16% per year. 4) The hedonic model was found to be a good predictor of actual market prices, despite the fact that it was originally estimated using list prices. Several variations of the model were examined, and, while the qualitative findings were robust, the precise estimates of the coefficients varied somewhat depending on the sample of products examined, the weighting of the observations and the functional form used in estimation, suggesting that the use of hedonic methods in this domain is subject to a number of limitations due, *inter alia*, to the potential for strategic pricing by vendors.
(*Network Externalities; Standards; Pricing; Software; Hedonic Regression; Information Technology*)

1. Research Problem and Model

1.1. Introduction

The production of packaged computer software for microcomputers is a multi-billion dollar industry in the US alone, and it is expected to be among the fastest growing industries over the next decade.[1] While a significant body of research examines the quality-adjusted prices of computer hardware, much less is known about software prices. The following analysis seeks to move forward in this area by building and empirically testing an econometric model identifying the factors affecting the price of an important category of microcomputer software: spreadsheets.

In microcomputer software, the purchase price reflects only a relatively small portion of the total consumer expenditure; learning and conversion costs account for much of the remainder. This, in turn, helps to create strong network externalities, with theory suggesting a consumer preference for software that is perceived as a standard (Farrell and Saloner 1985, Saloner 1989). We therefore focus on understanding and estimating the impact of installed base and standards on

[1] *Software Magazine* estimates that the top 100 independent software companies had combined revenues $13.9 billion in 1991 (Hodges 1992, p. 17).

0025-1909/96/4212/1627$01.25

the price of software. This effort requires developing a hedonic model that uses data not only on the features and prices of products, as is traditionally done when estimating price indexes, but also on the installed base of users. Insights gained in this area are likely to have direct implications not only for the software industry, but also for a variety of other technologies that exhibit similar economic behavior.

The organization of this paper is as follows. This section gives a brief overview of previous literature in this area and presents a description of the base model. The next section describes the data used to estimate the model. Results are presented in §3, along with additional analyses of the base model. Section 4 investigates several related research questions by examining additional specifications and subsamples. Section 5 then applies the model to an independent sample of market price data. Discussion and concluding remarks are presented in §6.

1.2. Network Externality Theory

There is a growing body of largely theoretical literature on the economics of network externalities and standards. The classic example of a product that exhibits network externalities is the telephone: having telephone service is only valuable if there are other people with compatible telephones that a user wishes to call. A more modern example might be electronic mail systems, which are valuable to the user only to the degree that many other people also use them. One of the contributions of this literature is an explication of the possible consumer benefits of standards, such as:

- network externalities from a community of users (e.g., ability to share information in a common format),
- a larger market for complementary goods and reduced market power of sellers,
- increased price competition, since competition on other dimensions is reduced and there may be decreases in cost due to production scale economies, and
- a thicker second-hand market, leading to a reduction in uncertainty (Farrell and Saloner 1985, Katz and Shapiro 1985, Westland 1992, Whang 1992).

However, these benefits are offset by certain costs of standardization:

- reduced product variety or diversity,
- "excess inertia" which can slow down adoption of better standards, and
- efficiency loss if the wrong standard is imposed (David 1985, Farrell and Saloner 1985).

A key implication of the theory is that the existing installed base of a technology affects consumer expectations and compatibility decisions (Farrell and Saloner 1986), although sponsorship of a technology by an industry leader can act to change expectations and/or reduce the importance of a pre-existing installed base (Cusumano et al. 1992, Fichman and Kemerer 1993). A limitation of this literature is that there has been little empirical validation of the theory. The practical relevance of the above insights depends on whether the network externality effects are of the same order of magnitude as other effects.

One approach to documenting the existence of network externalities empirically is to look for features which are designed to maintain compatibility. For instance, the choice of file format, user interface design or other interfaces is often influenced by a desire to create compatibility with as many users and products as possible. Greenstein (1993) finds that IBM compatibility is important in the mainframe market; users of IBM mainframes are more likely to buy mainframes from IBM in the future than are users of other vendors' products. Hartman (1989) uses a hedonic model with dummy variables to reveal that IBM compatibility is also important for microcomputers. In a paper closely related to this one, Gandal (1994) uses data from DataPro to construct a hedonic price index for spreadsheet software and determines that Lotus file compatibility has significant value.

In each case, the finding that compatibility with the dominant standard is valuable can be interpreted as evidence of network externalities, because the dominant standard has more users than any other standard. However, one could also interpret the coefficient on the compatibility variable in the same way as the coefficients on other feature variables—as a reflection of its intrinsic value[2]—or, alternatively, as an indication of the

[2] For instance, some users may consider the Lotus interface intrinsically easier to learn and use: "Quattro gives users the choice of selecting any of several optional interfaces. In general, our testers liked the Lotus-style screen and menus, rating 1-2-3 and the five "clone" programs highly for both ease of learning and ease of use" (National Software Testing Labs 1988, pp. 4–5).

"brand" value of being associated with a superior product.[3]

Classic network externalities are typically defined as an increase in the value of a product as the number of users of that product increases (Farrell and Saloner 1985, Farrell and Saloner 1986). Therefore, direct assessment of network externalities would examine the effects of the actual installed bases of competing products, including those which were not compatible with the dominant standard, on their prices.[4] This would provide a more direct estimate of the basic implication of network externalities: the price users will pay for a product is, in part, determined by the size of the network to which the product belongs. Software products in general, and spreadsheets in particular, make excellent test beds for this work. Learning time and compatibility are important for most types of software, giving rise to potential network externalities, and in the case of spreadsheets in the 1980s and early 1990s, there was a well-identified market standard, Lotus 1-2-3. Lotus 1-2-3 was touted as the single most successful computer application ever, and a key ingredient in the success of the highly popular IBM PC platform (Cringely 1992, p. 147).

1.3. The Hedonic Model

In order to estimate the magnitude of the effects of standards and network externalities it is important to control for other factors that can affect the price of software. We estimate a hedonic regression model using annual data on price and features for a set of microcomputer spreadsheet packages. Hedonic models are designed to estimate the value that different product aspects contribute to a consumer's utility or pleasure. According to Berndt (1991, p. 117), "implicit in the hedonic price framework is the assumption that . . . a particular commodity can be viewed as consisting of various . . . bundles . . . of a smaller number of characteristics or basic attributes. In brief, the hedonic hypothesis is that heterogeneous goods are aggregations of characteristics. Moreover, implicit marginal prices of the characteristics can be calculated as derivatives of the hedonic price equation with respect to levels of the characteristics."

This hedonic regression approach, developed in the 1920s, has been used on a wide variety of products and services ranging from asparagus to automobiles to marriage dowries (Rao 1993). A useful history and summary is provided by Berndt (1991). Perhaps the first successful application to information technology was in a study done by Chow, who estimated an annual quality-adjusted decline in mainframe computer prices of 21% from 1960–1965 (Chow 1967). More recently, Gordon (1993) found that the annual decline in prices was at least that large for computer prices through 1987. Berndt and Griliches use hedonic regression techniques to estimate the quality-adjusted change in prices for microcomputers for the period 1982–1988 (Berndt and Griliches 1990, Berndt 1991); they find that the decline in real prices averaged 28% per year.[5]

In order to estimate the optimal strategy for producers a number of researchers have focused on factors affecting hardware prices. An early paper by Michaels models the mainframe market, including such product feature attribute factors as size of main memory and the amount of secondary storage (Michaels 1979). Most recently, Rao and Lynch (1993) construct a hedonic model to estimate the value of attributes of computer workstations. Much less work has been done in the area of computer software. One notable exception is the work of Gandal (1994). He constructs a hedonic price index for spreadsheet software and finds that Lotus file compatibility has significant value, which provides evidence of network externalities. However, as Gandal notes, two important limitations of his research are the lack of data on unit sales of products and the lack of market price data.

This paper seeks to address these two limitations and extend the research on network externalities and the hedonics of software in a number of other ways. Most importantly, in order to provide a direct assessment of the

[3] The decision to make Microsoft's Windows operating system similar in style to the Macintosh operating system (and not DOS) presumably had little to do with a desire to tap into the relatively small installed base of Macintosh users and much to do with the widely-acknowledged superiority of a Mac-like user interface.

[4] Alternatively, one could look at the *potential* installed base, as Saloner and Shepard (1990) do.

[5] See also Triplett (1986) for an overview of the use of studies which apply hedonic methods to estimate computer price indexes.

role of network externalities, our research gathers data on unit sales of spreadsheet products to calculate the implicit price of having an installed base, as well as data on several alternative measures of Lotus compatibility. Because both installed base and the method used to measure Lotus compatibility turn out to be important, including these variables addresses a missing variables bias in the other coefficient estimates. Our analysis should therefore provide more accurate estimates of the values not only of network externalities and standards, but also of other features and the time trend. Furthermore, we seek to assess the reliability of the hedonic approach in this domain by attempting to predict market prices using our model with list prices. Additional extensions, including a comparison of new products with average practice, regressions based on a market-share weighted subsample, tests of the functional form of the model, and evaluation of the role of platform changes (e.g. from MS-DOS to graphical user interfaces) are also considered.

Accordingly, we construct a hedonic model to determine the implicit price of four general types of characteristics: 1) The effect of an increase in installed base for spreadsheet products is explicitly examined, thereby providing insight into the importance of network externalities. 2) The related, but distinct, question of the effect of standards on prices is also analyzed. 3) The more traditional, intrinsic feature variables typically used in hedonic price estimation are included. 4) A time trend variable allows us to discern any change in quality-adjusted prices due to technological progress or other factors.

Thus, our general model is:

$$P_{it} = f(N_{it}, S_{it}, F_{it}, T_t) \quad \text{where:} \tag{1}$$

P_{it} = Price, N_{it} = Network externalities attributes, S_{it} = Standards attributes, F_{it} = product Feature attributes, and T_t = Time trend, all of software package i (in year t).

The network externalities, standards and product feature attributes should positively influence price, whereas the expected sign on the time trend variable is negative, reflecting a decline in quality-adjusted price over time due to technical progress. We customize this general model to accommodate the appropriate hedonic variables for the microcomputer spreadsheet package market and then estimate it.

Table 1 Dependent Variable(s)

Variable Name	Definition
RLISTPRICE	Inflation-adjusted (real) list price for the product.
LOG_LIST	Natural log of the RLISTPRICE for the product.

2. Data

Given the model above, four kinds of data are needed for each product by year: 1) prices, 2) network externality variables such as installed base, 3) adherence to standards, and 4) product feature attributes to help control for intrinsic "quality." Much of the data used was originally compiled and described by MIT students Donna Mayo and Daniel Young (1993).

2.1. List Price Data

Reliable information on product pricing is clearly critical for this type of hedonic model. DataQuest and International Data Corporation (IDC) both very generously provided data on the spreadsheet market for 1987–1991.[6] For 1992, product list prices were collected from trade press reviews. As a result, the sample may under-represent the number of different products actually sold during 1992. The list price data included up to 22 unique products in each of six years; these were pooled to create a series consisting of 93 observations.[7] (See Table 1.)

Since these data include a time series component, the nominal prices require adjustment to account for inflation. Product prices were deflated to 1987 dollars using the GDP deflator, so the dependent variable is price in 1987 dollars.

2.2. Network Externalities

Packaged software exhibits positive network externalities in that the value of a product to an individual user increases to the degree that other people also use it (Ar-

[6] These two market analysis firms are the leading data sources for information on the software industry (Rebello et al. 1992).

[7] Data on market prices is discussed in §5.

thur 1988). Hence new users will prefer more popular spreadsheets to less popular ones, *ceteris paribus*. They will benefit from a greater abundance of third-party training opportunities and materials, complementary or compatible products, and user groups, and from an increased likelihood of vendor viability. It would thus be expected that products with a larger share of the installed base will exhibit a price premium over products with smaller shares.

The installed base of each product in each year was computed by summing its sales in all prior years, including sales of earlier, compatible versions.[8] This definition of installed base is consistent with industry parlance and with earlier work estimating the effects of installed base on the market share of minicomputers (Hartman and Teece 1990). The installed *base share*, in turn, was computed by dividing the installed base of each product by the sum of the installed bases for all products for a given year. This approach is likely to somewhat over-estimate the base share to the extent that there were other products in the market which were not in the sample, but this bias will be proportional for all products in a given year. In addition, according to DataQuest and IDC, these data represent the overwhelming majority of spreadsheet products sold.

2.3. Standards

In its pure form, a standard has no intrinsic value; rather, its value is derived by the adherence of other products to the standard. Given Lotus's pre-eminent market position, there should be an advantage to being Lotus-compatible. Since Lotus 1-2-3 had been the dominant product in the market since its introduction in 1983, many products have attempted to capitalize on Lotus' user base by providing the option to use, for example, an exact duplicate of the Lotus 1-2-3 menu tree. This copying of the menu tree was at the heart of the lawsuit between Lotus Development Corporation and Paperback Software International, which further indicates the perceived value of this particular standard.[9]

Table 2 Network Externality and Standards Variables

Variable Name	Definition
BASESHARE	Percentage of the then current installed base owned by this product.
LOT_MENU	Dummy variable, one if the product supports the Lotus 1-2-3 menu tree feature, zero otherwise.

Part of the value of the Lotus 1-2-3 menu tree stems from the greater ease of use for the installed base of users who already know the Lotus menu tree, as witnessed by this quotation from a contemporary software review:

> *Slash-F-R to retrieve a file. Slash-F-S to save. Slash-W-E-Y to clear the worksheet. Slash-C, mark the source range, mark the destination. Countless spreadsheet users are familiar with the command sequences popularized by Lotus 1-2-3 . . . Because many of these users are familiar with 1-2-3, the ideal program is one that uses the same command sequence.* (National Software Testing Labs 1990, p. 7)

The Lotus Menu variable also represents the source of the principal switching cost for Lotus 1-2-3 users. One can also examine the value of Lotus file compatibility, as Gandal (1994) did, or Lotus macro compatibility. However, of these measures, Lotus Menu compatibility is likely to be the most important, since a network benefit should be obtained via an opportunity to avoid the costs of incompatibility. A switch by an organization to another user interface would require an investment by every user in re-training, whereas incompatibility in reading and/or writing files could be solved through one-time construction or purchase of a converter. (Matutes and Regibeau 1988, Farrell and Saloner 1992). Therefore, a standards effect is represented by a Lotus Menu dummy variable, although variables for Lotus File compatibility and Lotus Macro compatibility will also be examined. The variable names are provided in Table 2.

2.4. Product Feature Attribute Data

Although we focus on the network externality effects, and specifically the value installed base share and a Lotus menu interface add to a product's price, other feature attributes are clearly part of a product's overall attractiveness and should be controlled for. It was desirable to find a single source that contained consistent and comprehensive information on product features. A

[8] No assumption was made to either depreciate or appreciate the installed base over time, since it was not obvious that early users were less (or more) influential than recent users.

[9] Lotus Development Corp. v. Paperback Software International, 740 F. Supp. 37 (D. Mass. 1990).

review of the comparative spreadsheet product reviews offered in major computer trade journals quickly revealed that the features reported on varied somewhat from year to year. Further, the set of products reviewed in any given year was often limited to only the most popular programs. Thus, it was not possible to rely entirely on reviews from any one journal.

One source that did meet the goals of the research was National Software Testing Lab's (NSTL) *Software Digest Ratings Reports* (National Software Testing Laboratories 1985). NSTL began publishing in 1984 and produced at least one and sometimes two issues dedicated to evaluating spreadsheet products in each year. Most of the products covered were business-class spreadsheets. (National Software Testing Laboratories 1988, 1988a, 1988b). Each NSTL report also contained detailed definitions of the product features. These definitions confirmed that the feature information collected was truly comparable across years for different products, which further increased our confidence in this source. The sample set covers products sold for the years 1987 through 1992. These data were supplemented and cross-checked with spreadsheet review articles from *PC Magazine*, *InfoWorld*, *Byte*, *Computerworld* and others. For instance, all articles on spreadsheet products since *PC Magazine's* inception in 1982 were examined. The data were also checked against product manuals, when available, for the more recent products.

Data were collected on the product feature attribute variables suggested as possible items that would influence purchase price decisions. The list of the variables and their definitions is presented in Table 3. A particular effort was made to collect the independent variables suggested by Gandal (1994) as significant predictors, and six of his seven "preferred model" variables are included.[10] In total, data on 28 independent variables were gathered including three alternative measures of Lotus-compatibility, four alternative operating systems, the four largest manufacturers, and 17 specific features.

[10] The only exception is Gandal's "WINDOW" variable, which takes the value of zero, one, or two depending on whether the maximum number of open windows is either 0, from two to fifteen, or sixteen or more. This information was not available for all of the observations.

2.5. Descriptive Statistics

The data sample was confined to commercial spreadsheet products. Neither shareware products, aimed primarily at home or casual users, nor extremely specialized financial modeling applications (e.g., Javelin and Encore) were included in this set. To allow the assessment of the price relationships more accurately, the data set includes only stand-alone spreadsheet programs, not spreadsheets sold as part of an integrated package. Integrated software packages generally include a word processor, a spreadsheet, a communications package, and other modules all in one box for one price. It was infeasible to determine the spreadsheet module's portion of the price of an integrated software package.

This data set contains consistent information for products representing at least 75% of units sold in each year, 1987 to 1991, according to DataQuest and IDC. Thus, it offers broad coverage of the spreadsheet market for products available for different computer platforms and operating systems during this period. With multiple years of data, both longitudinal and cross-sectional differences can be observed. These two independent sources were found to be relatively consistent with each other.

In total, the data set contains 93 different product observations, where a new observation is generated for each spreadsheet revision and each year that the revision is offered. It is worth noting that, despite the fact that a new version of a product may enter the market, sales of older versions continue. Practically, this results because new products are not neatly introduced at the start of a year, nor are older products withdrawn at a year's close. Also, there are users who choose to continue purchasing a perhaps out-of-date version of a product to maintain guaranteed compatibility with current files and applications or to avoid the hardware upgrades that new software releases sometimes require. This notion proved to be especially relevant for Lotus 1-2-3 products where, for example, sales of Lotus 1-2-3 Release 2.01 continued through 1991 despite its initial release being in 1985. Taking into account repetition of prior products, there are 68 distinct versions of products (such as Lotus 1-2-3 version 2.01 versus Lotus 1-2-3 version 3.0) in the data sample. The sample size of 93 is due to the fact that product versions appear in more than one year.

Table 3 Candidate Product Feature Attribute Variables

Variable Name	Definition (dummy variables equal one if the product has the feature, zero otherwise)
BACKCALC	One if the spreadsheet supports background recalculation.
CELLINKF	One if product can link spreadsheets by using a cell reference from an external worksheet in a formula in the current worksheet.
DB_COMPT	Equals the number of the following database file types that can be read or directly imported: dBase, Paradox, and SQL.
DDE	One if the product supports either the Microsoft Dynamic Data Exchange mechanism or the Mac Publish/Subscribe mechanism.
EMBEDCHT	One if the product has the ability to mix (embed) charts and spreadsheets on the same page.
EXT_DATA	One if the program provides links to external databases.
FONT_SUP	One if the product supports more than one font simultaneously on a worksheet or graph.
GRAPHS	One if product can create pie, bar and line graphs.
LAN_COMM	One if the product has the capability of linking independent user through a local area network (LAN).
LINKING	One if product can update values in multiple (linked) spreadsheets, zero otherwise.
LOGMINRC	Log of the minimum of the maximum number of rows and columns a product supports. A power variable.
LOT_FILE	One if product is file compatible with Lotus format. Equivalent to Gandal's "LOTCOMP".
LOT_MACR	One if the product supports Lotus macros.
MFR_BOR	One if the product is manufactured by Borland. A "make effect" variable.
MFR_CA	One if the product is manufactured by Computer Associates. A 'make effect' variable.
MFR_LOT	One if the product is manufactured by Lotus Development Corporation. A 'make effect' variable. Equivalent to Gandal's "LOTUS".
MFR_MS	One if the product is manufactured by Microsoft. A "make effect" variable.
MIN_CALC	One if the product supports minimal recalculation, a feature which enables the spreadsheet to recalculate only cells which need to be recalculated, rather than recalculating the entire spreadsheet when changes are made.
MOUSESUP	One if the spreadsheet supports at least three of the following mouse shortcuts: pulldown menus, drag & drop editing, speed formatting, speed filling, and icon/button bars.
OS_MAC	One if the product is designed to run under the Macintosh operating environment.
OS_OS2	One if the product is designed to run under the OS2 operating environment.
OS_WIN	One if the product is designed to run under the Windows operating environment.
PRNTPREV	One if the product can show an on-screen print preview.
SORT_COL	One if the product supports sorting by columns, in addition to the normal sorting by rows.
SRCH_RPL	One if the product supports global search and replace through cell contents.
WYSIWYG	One if the product supports a What You See Is What You Get interface.

Table 4 shows the distribution of the product set by year. The data set covers spreadsheets made by 11 vendors. Table 5 details how many data observations are attributable to products from the three major spreadsheet vendors. This data set includes 58 data observations for products that operate under MS-DOS on PC-compatible computer platforms and 35 data observations that run under graphical operating systems (MS-Windows, OS/2, and Macintosh) on Apple Macintosh or

Table 4 Distribution of Data Sample by Year

Year	Number of Data Points
1987	11
1988	15
1989	17
1990	19
1991	22
1992	9
TOTAL	93

Table 5 Distribution of Sample by Vendor

Vendor	Total Data Points
Lotus	23
Microsoft	21
Borland	9
Others	40
TOTAL	93

Table 6 **Distribution of Set by Operating System**

Operating System	Data Points
MS-DOS	58
Macintosh	17
Windows	12
OS/2	6
TOTAL	93

PC-compatible computers. Table 6 shows the distribution of the data set over these four operating systems, and Table 7 provides summary statistics for the data set.

There are several product feature attributes for which information was not readily available. For instance, while a spreadsheet's speed (e.g., file loading speed, recalculation speed, etc.) may enter into a buyer's purchase decision, obtaining comparable estimates of speed for many products is infeasible for a number of reasons. From the NSTL reports, speed ratings are available for only a small fraction of the data sample. The issue of product speed is further muddied by the fact that advances in microcomputer processor speed throughout the period under consideration make speed rates from year to year incomparable. Furthermore, it may also be true that the speed of spreadsheet operations is becoming a secondary issue simply because hardware advances can be counted upon to make up for deficiencies of the software itself.[11]

It would have been interesting to investigate product ratings for ease of use or learning, or other more subjective product qualities. However, such qualitative ratings from a consistent source are available for only a fraction of the data sample. In addition, more objective measures of product features may proxy for these more nebulous concepts. For instance, products that offer the WYSIWYG interface are often heralded for their ease of use. By quantifying the value of this interface, it may be possible to gain insight into the value created by producing a spreadsheet product that is easier to use than others in the marketplace.

Another concern may be that competitive pressures, hardware innovations, and other market factors also have an impact on spreadsheet software prices. We believe that the time trend variable may control some of these effects, with the rest being cap-

Table 7 **Descriptive Statistics**

Variable	Mean	Standard Deviation	Maximum	Minimum
Dependent Variables				
RLISTPRICE	338.36	135.46	615.73	59.85
LOG_LIST	5.70	0.57	6.42	4.09
Network Externality Variables				
BASESHARE	6.45	12.27	64.65	0.00
LOT_MENU	0.49	0.50	1.00	0.00
Candidate Product Feature attribute Variables				
BACKCALC	0.56	0.50	1.00	0.00
CELLINKF	0.56	0.50	1.00	0.00
DB_COMP	1.03	0.81	3.00	0.00
DDE	0.26	0.44	1.00	0.00
EMBEDCHT	0.46	0.50	1.00	0.00
EXT_DATA	0.76	0.43	1.00	0.00
FONT_SUP	0.89	0.31	1.00	0.00
GRAPHS	0.95	0.23	1.00	0.00
LAN_COMM	0.75	0.43	1.00	0.00
LINKING	0.74	0.44	1.00	0.00
LOGMINRC	6.32	1.74	10.40	5.54
LOT_FILE	0.98	0.15	1.00	0.00
LOT_MACR	0.62	0.49	1.00	0.00
MFGR_BOR	0.10	0.30	1.00	0.00
MFGR_CA	0.05	0.23	1.00	0.00
MFGR_LOT	0.25	0.43	1.00	0.00
MFGR_MS	0.20	0.41	1.00	0.00
MIN_CALC	0.69	0.47	1.00	0.00
MOUSESUP	0.27	0.45	1.00	0.00
OS_OS2	0.06	0.25	1.00	0.00
OS_MAC	0.18	0.39	1.00	0.00
OS_WIN	0.13	0.34	1.00	0.00
PRNTPREV	0.58	0.50	1.00	0.00
SORT_COL	0.45	0.50	1.00	0.00
SRCH_RPL	0.60	0.50	1.00	0.00
WYSIWYG	0.51	0.50	1.00	0.00

[11] For example, a contemporary review noted "To be sure, in this era of affordable 386-based machines, recalculation speed is a minor point compared with time spent constructing spreadsheets and preparing output to communicate their results" (Coffee 1991).

tured by the error term. Finally, all such estimations are limited by the fact that the data may be measured with error. This has been minimized in this study through the use of multiple contemporaneous industry sources.

3. Base Model Construction and Estimation

We propose the following general equation:

$$\text{LOG\_LIST} = \beta_0 + \beta_1 * \text{Installed Base} + \beta_2 * \text{Standard} + \beta_i * \text{Product Feature attribute } i + \beta_j * \text{Time} + \epsilon. \quad (2)$$

An initial, exploratory model was created that included all available network externality and product feature attributes. From this initial model a subset of the variables that were significant at the 90% confidence level (one-tailed test) were retained, and this smaller set formed the base model. As predicted, BASESHARE and LOT_MENU were among the significant variables. The other possible measures of Lotus compatibility, LOT_FILE and LOT_MACR, were not significant.[12] In addition, a relatively short list of five feature variables, pertaining to fairly distinct types of capabilities, also entered into the base model. The expectation is that these variables can collectively proxy for the overall "quality" of the product, including the effects of network externality features and time. The resulting *base model* equation is as follows:

$$\begin{aligned}\text{LOG\_LIST} &= \beta_0 + \beta_1 * \text{BASESHARE} + \beta_2 * \text{LOT\_MENU} \\ &\quad + \beta_3 * \text{EMBEDCHT} + \beta_4 * \text{LAN\_COMM} \\ &\quad + \beta_5 * \text{SORT\_COL} + \beta_6 * \text{SRCH\_RPL} \\ &\quad + \beta_7 * \text{WYSIWYG} + \beta_8 * \text{TIMETREND} + \epsilon. \quad (3)\end{aligned}$$

The dependent variable is the natural log of the list price of the product. The base case product, for which all the dummy variables were set equal to zero, is a commercial spreadsheet product that has no installed base of users; does not have a Lotus 1-2-3 style menu tree; cannot embed charts on the worksheet; does not support Local Area Network connections; cannot sort data by column; does not support a search and replace feature; does not have a WYSIWYG interface; and was sold in 1987.

The time trend variable is operationalized as TIMETREND = (Year-1987), and indicates the average decline in quality-adjusted price per year. The base model results are presented in the first columns of Table 8, labeled "Base."

The model was estimated using a heteroskedasticity-consistent covariance matrix as a control for possible heteroskedasticity in the sample (White 1980). The Belsley-Kuh-Welsch multicollinearity diagnostics indicated that the independent variables were not significantly confounded with each other (Belsley et al. 1980). Thus, their individual coefficients may be interpreted with relative confidence. The sensitivity of these results to a number of alternative specifications of the model are examined below. The signs of all the feature attribute variables were positive and were generally significant at the 99% level of confidence. The key exception to this is the SRCH_RPL variable, which, while significant at the 90% confidence level in the exploratory model with the complete set of variables, is not highly significant in the base model. However, this variable has been retained in the base model to avoid excessive fine-tuning.[13]

The constant term estimated in the regression may be interpreted as the predicted price (in logs) of a hypothetical base case product in 1987, including associated margins ($e^{4.87} = \$130.32$). The BASESHARE variable is positive and significant and, given the semi-log formulation of the model, can be roughly interpreted as indicating that a one percent increase in the share of the installed base of a product translates into slightly less than a one percent increase in the product's price.[14]

[12] When the base model was estimated with LOT_FILE or LOT_MACR instead of LOT_MENU, they remained insignificant, so their failure to enter into the model cannot be ascribed simply to possible collinearity.

[13] The results are little changed if this variable is omitted.

[14] Note that $e^{0.0075} = 1.00753$, so a one percent increase in the share of installed base is associated with a price that is 0.753% higher, ceteris paribus.

Table 8 Base Model and Specification Tests

	Base		Linear		Indiv. Years		Time Periods	
Variable	Coeff.	*t*-stat	Coeff.	*t*-stat	Coeff.	*t*-stat	Coeff.	*t*-stat
Intercept	4.87	*32.66*	132.50	*4.05*	4.75	*23.68*	4.88	*31.31*
Lot_Menu	0.38	*2.94*	102.40	*3.57*	0.38	*3.05*	0.36	*3.08*
Baseshare	0.0075	*2.64*	2.29	*2.97*	0.0073	*2.41*	0.0089	*2.59*
Embedcht	0.45	*3.37*	136.97	*4.14*	0.46	*3.34*	0.51	*3.66*
LAN_Comm	0.45	*3.64*	98.12	*3.61*	0.46	*3.65*	0.33	*2.11*
Sort_Col	0.33	*2.50*	91.35	*2.79*	0.32	*2.48*	0.51	*2.76*
Srch_Rpl	0.14	*1.23*	35.87	*1.42*	0.15	*1.40*	0.13	*1.21*
WYSIWYG	0.44	*3.67*	81.21	*3.30*	0.44	*3.76*	0.39	*3.10*
Timetrend	−0.16	*−4.09*	−39.16	*−4.58*			−0.18	*−3.18*
T88					0.07	*0.35*		
T89					−0.26	*−1.33*		
T90					−0.35	*−1.92*		
T91					−0.61	*−2.92*		
T92					−0.66	*−3.09*		
TLAN_Comm							0.27	*1.55*
TSort_Col							−0.28	*−1.52*
n	93		93		93		93	
R^2	0.56		0.58		0.58		0.58	
Adjusted R^2	0.52		0.54		0.51		0.53	
Durbin-Watson	2.00		1.84		2.05		2.00	
F-statistic	13.43		14.62		9.12		11.42	

This represents the average effect for the products in the sample over the time period 1987–1992. The LOT_MENU variable is also positive and significant, and exponentiating its value suggests that, on average, a product with the Lotus 1-2-3 menu tree interface commands a price approximately 46% ($e^{0.38} - 1$) greater, all else being equal.

The set of product feature attribute variables can be interpreted similarly. As the primary emphasis of this analysis is on the network externality features, the product feature attribute variables are of interest mainly for examining their relative magnitudes versus the base share and Lotus standard variables. In general, these variables are approximately the same order of magnitude as the Lotus menu variable, ranging from 0.33–0.45. This suggests that network externality-type variables play an important role in the determination of price.

The time trend variable shows a decline over the period studied. This may be interpreted as the price decrease in each year, holding the quality factors constant. Over the entire period, the results imply that quality-adjusted price declined in real terms by an average of approximately 16% per year, which is in contrast to the slight increase in the nominal price of these products over this period. However, as shown below, the rate of the decline is sensitive to the choice of observations included in the sample and how they are weighted.

3.1. Specification Tests

As noted above, the network externality and feature variables were selected because they were believed to represent important product feature attributes and because of data availability. One issue with such a hedonic model is whether other critical feature attribute variables have been omitted. As one way of testing whether the model is incorrectly specified, Berndt and Griliches (1990) argue that any quality change not captured by the selected feature attribute variables should be unrelated to the "vintage" of the product, i.e. the first year the product was available, or a product's age. This test involves adding a set of vintage variables for each of the years in the sample, as well as a set of age variables representing the number of years a product has been available. An F-test is used to compare the unrestricted model, which contains both the vintage and age variables, to a restricted model without these variables. The restricted model passes the F-test at the 5% confidence level (F-stat $= 1.59 < f(11, 93 - 20) = 1.91$). Similarly, a test of the restricted (basic) model shown in Table 8 versus an unrestricted model with age dummies also passes the F-test at the 5% level (F-stat $= 0.36 < f(6, 93 - 15) = 2.18$).[15] Passing these specification tests is consistent with the hypothesis that the choice of hedonic feature variables captures the bulk of the quality differences among products.

3.2. Functional Form of the Model

The base model was estimated using a semi-log form of the regression equation, which is common in the literature (e.g., Michaels 1979, Berndt and Griliches 1990, Gandal 1994). A justification for using log prices as the dependent variable is that each quality improvement is like getting a larger "package" of the good at the same price, so that the effect on consumer utility is multiplicative (Fisher and Shell 1971). Although this argument seems pertinent to the case of spreadsheet software, the hedonic function in general is the envelope of many disparate supply and demand functions (Rosen 1974, Epple 1987) and, as argued by Triplett (1989), other functional forms may also be appropriate. In particular, an alternative model would be a pure linear form that expresses the dependent variable, product price, in dollars, such as that applied by Rao and Lynch (1993). (The choice of functional form does not affect the inclusion of any of the dummy variables.) Berndt, Showalter and Wooldridge (1990) found that the choice of functional form could, in some cases, make a significant difference in the calculation of price indexes. Accordingly, as a test of the robustness of the results, the base model was also estimated using the linear form, and these results are presented as the Linear model in Table 8.

[15] Details of these tests, including all coefficient values and t-statistics are available in Brynjolfsson and Kemerer (1994).

When the Linear model is compared to the Base model in Table 8, it is apparent that the significance of each of the variables is very similar, as indicated by the t-statistics. Furthermore, the coefficients in each of the semi-log and linear regressions represent comparable fractions of the average product price, e.g., the value of $2.29 for each point of BASESHARE corresponds to an increase of 0.0068 percent (2.29/338.36, the average price), compared to a value of 0.0075 from the semi-log model. This suggests that the Base model findings are not an artifact of the semi-log specification.

3.3. Individual Year Time Dummy Variables

In the Base model the residual effects captured by the passage of time are reflected in a single dummy variable, TIMETREND. While preserving degrees of freedom, this approach does constrain the change in prices over time to be equal in each of the years. An alternative form of the model includes five separate time dummy variables (e.g., T88, T89 . . . T92) in place of the single TIMETREND variable. This model was also estimated, and the results were similar to those of the Base model, as shown in Table 8. The time dummies showed a fairly steady pattern of decline, and the other coefficient values in the model remained essentially unchanged. An F-test of the two models did not reject the null hypothesis of a linear time trend at the 5% level (F-stat $= 0.778 < F(4, 93 - 13) = 2.53$ @ 0.05). Therefore, the more parsimonious Base model is preferred. Figure 1 shows the average change in price by year for the model with individual time dummies.

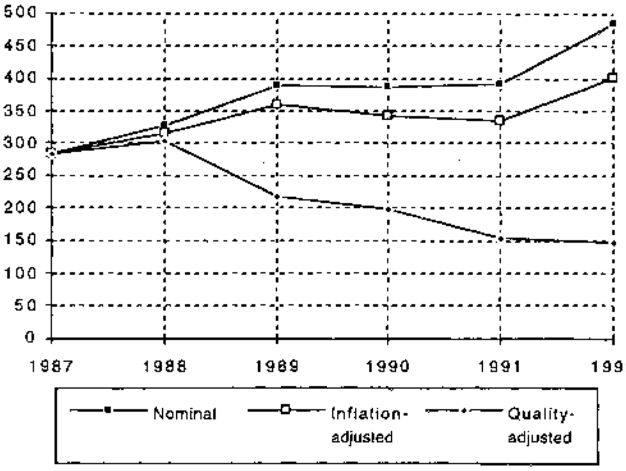
While a portion of the quality-adjusted price change between any two years may be due to changes in the sample composition, the overall time pattern of quality-adjusted prices suggests that consumers are clearly better off. Given the large research budgets of spreadsheet vendors, some of this improvement is presumably due to continuous innovation in spreadsheet technology.

Two other factors, however, may be important. First, it is not uncommon for software features to be limited by the availability of sufficiently powerful computer hardware. In particular, hardware is a complement and potential constraint for spreadsheet software features. The well-documented decline in the price/performance ratio of computers could enable the provision of more advanced software features, such as a WYSIWYG display, which would have been unattractive or even infeasible when the typical processor speed was slower and memory was lower. Thus, advances in computer hardware may spill over into software quality (Gurbaxani and Mendelson 1990). Second, the 1988–1992 period was a period of increasing competition in the spreadsheet market as Lotus's dominance was challenged by Borland and Microsoft, and this may have lowered margins and therefore prices, transferring surplus from producers to consumers. Because all three of these interpretations (technical progress, declining hardware prices, and increased competition) imply similar effects on all products in a given year, they cannot be further distinguished given these data.

3.4. Stability of the Base Model over Time

In estimating a model on data that are both cross-sectional and longitudinal, one concern is whether the model is stable over time, i.e., whether the model fits equally well at the beginning of the period as at the end. Therefore, as a test of the stability of the parameters over time, an extended model was created. This extended model allowed all the parameters to vary in each of the two periods through the addition of a set of time-period ("*T*") interaction variables (e.g., TBASESHARE) which were the product of the original variable value and either 0, if the observation was for the years 1987 through 1989, or 1, if the observation was for the years 1990–1992. A significant positive (negative) value for a "*T*" variable would indicate that the feature was relatively more (less) valuable in the later period. When this model was estimated on the data set there was no significant explanatory power added to the base model; an *F*-test did not reject the null hypothesis that all the parameters were stable (F-stat $= 0.98 < f(9, 93 - 18) = 2.04$ @ 0.05). To help guard against the possibility that the power of the test was simply insufficient to reject the null, a second, more focused test was conducted. An inspection of the individual parameters indicated that TLAN_COMM and TSORT_COL might be exceptions to the general finding that all of the parameters appeared stable. Therefore, an additional model was created, consisting of the base model plus just these two additional variables. The results of this model, shown in Table 8, also do not allow rejection of the maintained hypothesis of stable parameter values over the two periods (F-stat $= 2.04 < f(2, 93 - 11) = 3.15$ @ 0.05), and further, the individual *t*-statistics for the two added variables were not significant at usual levels.

Figure 1 **Average List Price ($), 1987–1992**



These results suggest that the base model's estimates are relatively stable over the time period studied. However, it is worth noting that the base model parameter estimates only represent average values over the entire period, and that a larger data set would permit a more fine-grained analysis by individual year.

4. Further Analysis of Network Externalities

The previous section presented the base model and analyzed several important basic modeling decisions. This section extends the analysis by examining six variants on the Base model.

4.1. Sensitivity Analysis of Network Externality Variables

In the base model both of the network externality variables, BASESHARE and LOT_MENU, are significant explanators of price. What is the effect on the model if one or the other of these variables is not included? In order to examine this, the Base model was rerun twice, dropping each of these variables in turn, and the results are shown as the first two models in Table 9. In each case the coefficient on the remaining variable increased (for LOT_MENU from 0.38 to 0.42, and for BASESHARE from 0.0075 to 0.0099). This suggests that omit-

Table 9 Sensitivity Analysis

	w/o Baseshare		w/o Lot_Menu		WLS		New Products		Comb. Baseshare		Make Effect		DOS Only	
Variable	Coeff.	*t*-stat	Coeff.	*t*-stat	Coeff.	*t*-stat	Coeff.	*t*-stat.	Coeff.	*t*-stat	Coeff.	*t*-stat	Coeff.	*t*-stat
Intercept	4.89	*33.21*	5.12	*41.31*	4.70	*34.16*	4.75	*26.42*	4.86	*33.26*	4.87	*29.40*	4.48	*23.66*
Lot_Menu	0.42	*3.46*			0.56	*4.47*	0.58	*3.82*	0.37	*2.49*	0.37	*2.33*	0.76	*4.85*
Baseshare			0.0099	*3.59*	0.008	*3.16*					0.0071	*2.34*	0.0091	*2.17*
Embedcht	0.43	*3.19*	0.39	*2.60*	0.27	*2.00*	0.67	*3.63*	0.43	*3.18*	0.45	*3.36*	0.38	*1.48*
LAN_Comm	0.52	*4.34*	0.44	*3.00*	0.47	*3.63*	0.45	*3.23*	0.54	*4.75*	0.45	*3.39*	0.65	*2.96*
Sort_Col	0.27	*2.11*	0.11	*0.98*	0.57	*4.00*	0.64	*4.19*	0.27	*2.11*	0.33	*2.49*	0.52	*2.13*
Srch_Rpl	0.13	*1.15*	0.28	*2.80*	0.19	*1.42*	0.09	*0.62*	0.14	*1.24*	0.14	*1.22*	0.19	*1.14*
WYSIWYG	0.45	*3.82*	0.31	*2.47*	0.16	*1.72*	0.41	*2.86*	0.45	*3.78*	0.44	*3.32*	0.19	*1.52*
Timetrend	−0.16	*−4.19*	−0.15	*−3.71*	−0.09	*−2.16*	−0.24	*−4.78*	−0.16	*−4.29*	−0.16	*−4.08*	−0.20	*−4.09*
Combaseshr									0.001	*0.52*				
Mfgr_Lot											0.03	*0.22*		
n	93		93		74		41		93		93		58	
R^2	0.54		0.50		0.33		0.70		0.54		0.56		0.63	
Adjusted R^2	0.50		0.46		0.25		0.64		0.50		0.51		0.57	
Durbin-Watson	2.06		1.97		1.76		1.16		2.10		2.01		1.70	
F-statistic	14.34		12.34		14.35		11.11		12.48		11.80		10.59	

ting BASESHARE from the regression would lead to slightly biased estimates of LOT_MENU and vice-versa.

4.2. Weighted Least Squares (WLS) Estimation

Not all of the products in the sample were equally successful in the market. Those products with smaller market shares may have been mis-priced by their vendors or may have been purchased primarily by a market niche of customers instead of the mainstream. Although each of the products represents purchases of some consumers and, in equilibrium, should reflect the underlying hedonic prices of the characteristics they embody, it might be argued that it makes sense to weight more heavily the products with larger market shares. In particular, if the variance of errors in observed price is inversely proportional to a product's market share, then weighting each observation by the square-root of market share before running OLS is the appropriate correction for the resulting heteroskedasticity. Using data on the units sold for each product each year, market share was estimated for 74 of the products in the sample.[16] This estimate of market share is based on dividing the units sold for each product by the sum of all units sold in a given year. The results of this regression are given as the third model in Table 9. As compared with the base regression, the values of BASESHARE and LOT_MENU are both slightly higher, indicating that network externalities may be even more important when successful mainstream products are given more weight. The relative values of the feature attribute variables change somewhat, but the most notable difference is the decline in TIMETREND which from −16% to approximately −9%. Apparently, much of the improvement in the average quality-adjusted price in the sample is due to improvement in products with relatively smaller market shares, so consumers of the most popular products are less likely to have benefitted.

4.3. Only New Products

More encouraging news for consumers is provided by a regression on new products alone. Forty-one products in the sample were not sold in any previous years.[17] To run the regression on this subsample, the variable BASESHARE was dropped, since these products

[16] Unit-sales data were not available for 1992 and for a small set of products in other years.

[17] Including major version changes and platform changes.

obviously tend to have no installed base. As shown in Table 9, new products, which might be construed to be at the technological "frontier," appear to experience a more rapid quality-adjusted price decline over time, nearly 24% (versus 16% for the full sample). Berndt and Griliches (1990) found a similar phenomenon in the microcomputer market, where they report that a "new model only" price index declines more rapidly than indexes based on the full-sample. When compared to the weighted least squares estimation, which captures the buying habits of the bulk of users, this regression suggests that the technological frontier grew more distant from average practice over time. One interpretation of this finding is that inertia increased in the spreadsheet market as users passed over innovative products in favor of products with larger installed bases. Interestingly, the coefficient for LOT_MENU for this subsample is 0.58, vs. 0.42 in the model run on the full sample without BASESHARE.[18] It appears that the Lotus menu tree interface may have been even more valuable in the sample of new products than in the full sample, perhaps because these products disproportionately needed to "piggy-back" on the existing installed base of users of the industry-standard interface.

4.4. Measurement of Installed Base Share

In the Base model, BASESHARE is measured as the percentage of the installed base that is owned by a particular product; e.g., the BASESHARE of Lotus 1-2-3 is the percentage of the installed base that were sales of Lotus 1-2-3. This is consistent with how market share is traditionally defined. An alternative definition, however, considers installed base as the share owned by all menu-compatible products. For example, the BASESHARE for Lotus 1-2-3 would not simply be Lotus 1-2-3 installations, but Lotus 1-2-3 and all compatible products with menu tree interface. This new variable, COMBASHR (for COMbined Baseshare), was computed for all the data and substituted for BASESHARE in the base model regression.

The results are shown as the fifth model in Table 9, and are generally a poorer fit to the model. Adjusted R^2 drops slightly from 0.52 to 0.50, but more importantly, the t-statistic on COMBASHR is 0.51, versus 2.49 for BASESHARE. (Nor does COMBASHR add explanatory power to the Base model when it is included in addition to BASESHARE.)

One possible explanation for the fact that COMBASHR is not as good an explanator of price is that COMBASHR simultaneous includes elements that are associated both with increases and decreases in price. The network externality effect, that is, a larger installed base of compatible products, would tend to make the product more valuable to consumers, exerting upward pressure on the price. On the other hand, an increase in COMBASHR can also reflect a larger installed base of a different vendor's compatible product, which is, by definition, a relatively close substitute. This would tend to increase the strength of competition and exert downward pressure on prices. Therefore, the failure of COMBASHR to be a significant predictor suggests that these two effects roughly cancel each other out.[19]

4.5. Make Effect

In addition to the specific network externality benefit of the Lotus menu tree structure, there may be other benefits that consumers ascribe to Lotus products. Given Lotus's early and dominant market position, they were seen as a market leader during this period. Therefore, their long-term viability was unlikely to be a concern, and a consumer making a purchase decision could assume that ongoing vendor support would be assured. In addition, continued improvements to the product in the form of new releases could be expected. Finally, the Lotus brand name may have acted as a proxy for a general product quality variable not captured by the other variables. Collectively, these effects are captured by the Lotus manufacturer dummy variable, or "make effect" (Berndt 1991).

An alternative version of the model was tested adding a dummy variable equal to one if "manufacturer = Lotus Development Corporation," and zero otherwise (MFR_LOT). As shown as the "Make Effect" model in Table 9, the addition of this variable was not

[18] Note that this is the appropriate comparison, and not to the value of 0.38 in Table 8, as that result is based on two model differences, use of the full sample and inclusion of the BASESHARE variable.

[19] A variable OTHBASHR, defined as (COMBASHR-BASESHARE), was also examined. When this variable was substituted for BASESHARE, its sign was negative, but not statistically significant.

significant (t-statistic = 0.22) while LOT_MENU remains significant (t-statistic = 2.33) despite some collinearity between MFR_LOT and LOT_MENU. Nor are the other feature attribute variables and the time trend significantly affected. This suggests that the variables in the base model, including LOT_MENU, capture most of the perceived quality associated with Lotus products.

If, instead of adding the MFR_LOT variable to the base model, it is substituted for LOT_MENU, the fit of the model does not improve, as might be expected if LOT_MENU were simply proxying for part of an overall make effect. Instead, the R^2 declines slightly to .52. This supports the argument that the provision of the Lotus menu capability is a good predictor of price, and, as tested on this sample, a better predictor than the make effect.

4.6. Change in Technological Architecture

An issue of increasing importance in technological systems is a product's compatibility with complementary product standards. In the case of packaged software, these applications must be compatible with an existing architecture, or platform, consisting of computer hardware and operating systems.

Compatibility with complementary technologies is considered an important determinant of a product's success, and therefore a change in the complementary technology offers an opportunity for new entrants (Farrell and Saloner 1987). Recently, several authors have suggested that such shifts occur in the microcomputer software market; for example, Cringely notes the opportunity for Lotus 1-2-3 that was created by the development of the IBM PC/DOS platform as an alternative to the Apple II platform, where the Lotus 1-2-3 predecessor, VisiCalc, was already established (Cringely, 1992). Morris and Ferguson suggest that a similar change took place as the IBM PC/DOS platform that was Lotus 1-2-3's home was replaced by graphically oriented "architectures" centered on Windows which provide an opportunity for Microsoft's Excel product to become dominant (Morris and Ferguson 1993).

Therefore, in order to examine this effect, the dataset was split into two subsets, one consisting of products operating under the older, dominant operating system, DOS, and the other consisting of the non-DOS (i.e., Mac, Windows, and OS/2) products. Collinearity prevented the full model from being estimated separately on each data set: for the non-DOS products, the value of the WYSIWYG variable was 1 for every observation in the sub-sample. Instead the model was estimated on the DOS subset ($n = 58$), and a predictive Chow test of the Base model and the DOS-only model did not reject the null hypothesis of equality of parameters (F-stat $= 0.54 < f(34, 58 - 9) = 1.65$ @ 0.05). This suggests that it may be appropriate to estimate the model as the "microcomputer spreadsheet market" rather than more narrowly as, for example, the "DOS spreadsheet market." However, it is worth noting that the LOT_MENU coefficient estimate from the base model (0.38) for the entire sample is less than that for the DOS-only subsample (0.76), which is consistent with what would be predicted by theory, and with what has been suggested by Morris and Ferguson.[20]

4.7. Sensitivity Analysis Summary

Taken as a whole, the combined results of the Berndt-Griliches specification tests, the alternative functional forms, the weighted least squares regression, and the various alternative regressions provide some indication of the relative degree of variability of the parameter estimates. In each case the results were qualitatively similar: the null hypothesis of no network externalities was rejected by the coefficients on both BASESHARE and LOT_MENU, the product feature attribute variables were collectively important, and the TIMETREND was negative. However, the specific values of the variables, as described above, did vary. This could be due to heterogeneity in the sample, measurement errors in the data, and/or disequilibrium in the spreadsheet market. Accordingly, the specific values of the coefficients from regressions such as those in this paper should be interpreted with appropriate caution.

[20] Similarly, a regression which included MFR_LOT in lieu of LOT_MENU showed a decline in the Lotus make effect on the non-DOS platforms, which may indicate that Lotus earned less of a premium when its menu became less of a standard, or, possibly, that its non-DOS products were perceived to be of lower "quality" in dimensions which were not otherwise captured by the base model.

5. Market Price Validation

5.1. Market Price Data

The IDC and DataQuest data reflect a package's list price. This is consistent with hedonic pricing models used in previous computer hardware and software studies, (e.g. Cole et al. 1986). However, because much microcomputer software is sold at a discount from list price, using list price data to develop the hedonic model may introduce bias if discounts across products are systematically related to the explanatory variables in the model.[21] The large gross margins for most software products make this a very real possibility, especially with regard to the variables for network externalities and standards (i.e. BASESHARE and LOT_MENU). Therefore, results from list price data alone may not be sufficient to conclude that there is an effect on actual market prices. In order to check whether a bias resulted from using list prices, a third, independent set of price data was constructed from the prices of the software packages as advertised in major microcomputer trade magazine advertisements. Market prices were found for a subset of 55 of the 93 observations in the full sample. The average discount was found to be 30.5% off list price and a correlation of 0.88 was found between market and list prices.

5.2. Validation of the Base Model on Market Prices

The results for the subset of 55 market price data observations, presented in Table 10, are very similar to the results of the Base Model using list prices. The coefficients on BASESHARE and LOT_MENU are little changed, while, among the feature attribute variables, SORT_COL loses some significance, and SRCH_RPL gains some in the market price model. Interestingly, the fit of the entire model, as measured by adjusted R^2, has improved. A Chow-test was run of the null hypothesis that the coefficients (except for the constant term) were identical for list price and market price. This test was unable to reject the null hypothesis (F-stat $= 0.33 < f(8, 55 + 55 - 16) = 2.02$ @ 0.05).[22]

Table 10 Ln of Real Market Price is the Dependent variable ($n \approx 55$)

Variable	Coefficient	t-stat
C	4.14	19.16
BASESHARE	0.0073	2.24
LOT_MENU	0.38	2.10
EMBEDCHT	0.57	2.59
LAN_COMM	0.79	3.84
SORT_COL	0.34	1.63
SRCH_RPL	0.34	2.20
WYSIWYG	0.47	3.15
TIMETREND	−0.22	−3.81
R^2	0.68	
Adjusted R^2	0.63	
Durbin-Watson	1.39	
F-statistic	12.47	

In comparing the estimates shown in Table 10 with the Base model shown in Table 8, two aspects of the data have changed. The first is the values of the dependent (price) variable (Table 10 uses market prices, while Table 8 uses list price), and the second is that the market data set represents a subset of the products of the list price data set (Table 10 results are based on an n of 55, vs. the n of 93 in Table 8). To examine whether the latter was significant in understanding the improvement in fit, the model was re-estimated on a matched sample of the list price data set consisting of only those product observations that also appeared in the market price data set. This model has an R^2 of 0.64, better than the base model on all 93 observations, but still not quite as good as that achieved on the market price data.

Figure 2, based on a model with five separate time dummies instead a single linear time trend, shows a

[21] In addition, an increasing portion of sales are also accounted for by upgrades, generally at much lower prices than list. Because the value of upgrades is a function of the *difference* in the features between the upgrade and the existing product, it is not appropriate to include upgrade prices in the same model as full-priced products to new users, and so, consistent with prior research, the sample does not include any upgrade-only prices.

[22] Interestingly, Berndt and Griliches, in their hedonic model of microcomputer hardware prices, also do not reject the notion of parameter equality (aside from a parallel shift) in the list and market data sets in perhaps the only previous hedonic study of information technology that uses a combined data set of both list and market observations, (28% of their hardware price observations were from market price data) (Berndt and Griliches 1990).

similarity between the pattern of the decline in quality-adjusted market price of spreadsheet software and the list price data results shown in Figure 1, with the effect being somewhat more dramatic in the market data.

5.3. Predicting Market Prices

The results of the previous section suggest that the base model may be a useful predictor of market prices. To examine this idea, a simple new model was created with LOG_MRKT, the log of the market price as the dependent variable, and $\hat{Y}_{\text{list}}$, the *predicted* value of the list price obtained as a result of running the base model, as the independent variable.

The results of this model are (*t*-statistics are given in parentheses):

$$\text{LOG\_MRKT} = \underset{(-3.12)}{-2.24} + \underset{(10.59)}{1.30}(\hat{Y}_{\text{list}}),$$

$$R^2 = 0.65, \quad F\text{-stat} = 98.06.$$

For comparison, the original model was rerun using the log of list price as the dependent variable for the matched set of 55 observations:[23]

$$\text{LOG\_LIST} = \underset{(-1.32)}{-0.84} + \underset{(10.63)}{1.14}(\hat{Y}_{\text{list}}),$$

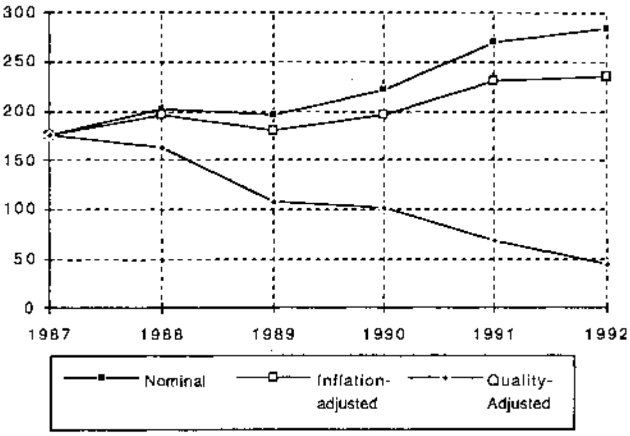
$$R^2 = 0.63, \quad F\text{-stat} = 89.39.$$

Strikingly, a comparison of the R^2 of each model indicates that the coefficients from the base model are slightly better predictors of market price than they are of the list price data from which they were estimated. This is consistent with the notion that the market price data reflect the true value placed on features by consumers better than the list price data do, and suggests that the results of the base model can be used to predict the eventual market price.

Further confirmation of this was provided by an expanded version of this model with an additional independent variable, LOG_LIST, the natural log of the list price. LOG_LIST would be expected to be a strong predictor of LOG_MRKT, the natural log of the market price. The results of this model are:

[23] Note that running this model on all 93 observations would give a value of 0 for the intercept and 1 for the other coefficient.

Figure 2 Average Market Price ($), 1987–1992



$$\text{LOG\_MRKT} = \underset{(-4.68)}{-1.51} + \underset{(9.92)}{0.89}(\text{LOG\_LIST}) + \underset{(2.60)}{0.29}(\hat{Y}_{\text{list}}),$$

$$R^2 = 0.88, \quad F\text{-stat} = 193.35.$$

Even with the inclusion of the actual list price data, the model results still add statistically significant predictive power. This can be interpreted as evidence that the hedonic model conveys information about the market value of features above and beyond the information implicit in actual list price decisions. If two products have the same list price, the market price will tend to be higher for the one with higher quality, as measured by the Base model.

6. Discussion and Future Research

6.1. Interpretation of the Findings

The prices of spreadsheets are potentially determined by a large number of factors other than network externalities, feature attributes, and a time trend. Technical progress and turbulence in industry structure suggests that prices are not in complete equilibrium. Furthermore, because of the relatively low marginal cost of producing spreadsheet software, vendors have considerable discretion in setting the list prices for their products. As a result, strategic pricing and other factors beyond the scope of the hedonic model may be significant, so that estimating precise hedonic functions is likely to be difficult. Nonetheless, the hedonic hypothesis—that the price of a software package can be modeled as a function of its features—is ultimately an empirical question. The analysis above provides support for this hypothesis, indicating that such a model can be used to identify characteristics which are important and to estimate the relative values of these characteristics. In particular, our model suggests that the positive network externality effects from installed base and from compatibility with a dominant interface standard are approximately as important as any of the intrinsic product features. These results provide additional empirical data for the ongoing economic study of innovations and compatibility. However, because the exact values of the coefficients varied depending on which observations were included and how they were weighted, so precise quantitative conclusions about the values of specific features should be made with care.

These results also have a number of practical implications. For software vendors they offer insights into appropriate package offerings. Developers can get a sense of the marginal value attributable to either features previously provided or currently provided by other vendors in order to glean insights about relative consumer preferences. Marketing staff can infer lessons about the relative rate of price change over time and therefore the likely duration of any existing market premiums. Legal staffs can use these results to estimate the market value of compatibility with an existing standard for licensing purposes. Another implication of the results is that the prices of spreadsheet software, adjusted for features, have declined significantly, so consumers are receiving significantly more sophisticated products at the same cost. The 16% price decline found in the Base model, while somewhat less than the quality-adjusted price decline in hardware, contrasts with the widely-held notion that software development has shown little progress.

6.2. Competitive Strategy

The evidence that network externalities are important in this market has significant implications for competitive strategy. From the perspective of the vendor, it is advantageous to achieve a significant share of the market quickly, so that higher prices can be commanded due to the positive effects of consumers joining a large network. This suggests that the initial purchase should be made as easy as possible, understanding that the early adopters can be seen as, to a degree, "subsidizing" later adopters, given that the sum of intrinsic and network benefits is lower for early adopters than for later adopters.[24]

In addition, vendors who seek to control important standards should consider licensing their technology, not only for the revenue stream such licensing offers, but also to increase the size of the network and therefore increase their own future sales. While licensing creates potential competitors, the sensitivity analysis using COMBASHR suggested that in the spreadsheet market the network externality effect may balance out the increased competition, so that the net effect on the vendor's own products may not be as negative as they would be for products which had no positive network externalities. In addition, by licensing the technology the vendor avoids creating significant incentives for other vendors to innovate around the standard, which could produce important competing products with distinct user bases.

Vendors should also consider that changes in the value of network externality attributes over time may be related to changes in complementary goods, specifically the hardware and operating system platforms that the application software requires to run. Entry into the software market is more likely to be successful when there is a major change in platform because of inertia from network benefits. However, when there is a significant change in the architecture, this may tip the balance and make users willing to abandon their previous application software. For example, the evidence that the Lotus menu tree interface is less valuable under a

[24] In fact, the correlation found between higher installed base and higher prices may be due, in part, to the realization by vendors that earlier adopters should be charged a lower price in order to build the network. An extreme example of this in the personal financial management software market was the decision by Computer Associates to "give away" one million copies of their product, *Simply Money* for a "shipping and handling fee" of $6.95 in order to build an installed base. Application of a hedonic model similar to the one developed here to that market segment might have provided some indication of whether the value of resulting installed base was likely to be worth the investment.

graphical user interface supports the argument that the initial absence of a Lotus-based product allowed entry and significant network creation by Microsoft Excel. This reduced the benefits to new users of adopting the Lotus standard, since the prevailing standard, in the sense of widest use, was not Lotus-based. Indeed, while Farrell and Saloner model a constant arrival rate of new customers, they speculate that a non-constant arrival rate will lead to more entry during high growth periods (Farrell and Saloner 1986). Another way to look at it is that a new hardware or operating system platform "levels the playing field" and erases the advantage of the incumbent network.

6.3. Future Research

A direct implication of these findings is that installed base can be treated on a par with intrinsic product quality in affecting the market value of spreadsheets. Because a product cannot build an installed base until it is released, this suggests that product managers face a direct and quantifiable trade-off between releasing a product later (but with more features) versus releasing it earlier. Further research could examine the strategic behavior of software companies to determine the extent to which they "invest" in installed base by rushing development, lowering margins, increasing advertising, or foregoing features which take time to implement. Different firms may choose different strategies depending on the existence of complementary products in their portfolio or constraints on their finances.[25]

The role of complementary products in supporting or undermining network externalities is likely to be increasingly important in the information technology field. The information technology industry as a whole is evolving toward the prevailing model of the personal computer industry, in which different vendors control hardware, operating systems and various layers of applications software.

An obvious extension of this work would be to examine more closely the role of platform changes, such as the change from DOS to Windows, on the value of standards and network externalities. If, as expected, it is found that success rates are markedly better for entrants at certain opportune times, it would obviously be of great interest to the business community. Furthermore, the relative values of coordinating product features and introduction by a single vendor (as with software "suites") versus alliances of several vendors versus industry-wide standards will depend, *inter alia*, on the size of network externalities. Implications for government policy also could be drawn from such findings. For instance, the evidence found here of user inertia, manifested as a growing gap between average practice and the technological frontier of new products, may be socially inefficient.

Other, shorter term future research would involve extending this model into future time periods and/or adapting the model developed here to other categories of microcomputer software. For example, word processing software has had a number of important vendors and products (WordStar, WordPerfect, Microsoft Word), as opposed to the historical near-monopoly position of Lotus in the business spreadsheet market. In the database market, Ashton-Tate's early dominance with dBase has lessened without a single dominant newcomer. It might be expected that results for these markets would differ from spreadsheets.

In addition, with such data on other categories it would be possible to develop a richer model of a greater portion of the total microcomputer software market by estimating a system of simultaneous equations that would allow for various cross-equation restrictions. For example, vendors with products in multiple categories could be expected to benefit from this breadth of offering through a stronger make effect. This make effect variable could be restricted to be equal across equations if that was felt to appropriately represent the realities of the market. Alternatively, the degree to which reputations varied across product categories could be tested. Similarly, compatibility effects may also differ in, for example, the case where no single standard dominates.

Finally, the hedonic model applied in this paper, while widely used, can be expected to underestimate price declines to the extent that it does not incorporate the value of new features which are introduced over time. Recent work by Bresnahan, Berry and Trajtenberg

[25] In this spirit, Chevalier and Scharfstein (1995) present evidence that liquidity-constrained firms in the supermarket industry are less able to invest in building market share than are firms with greater access to capital.

has resulted in a methodology for more accurately assessing the benefits of new products and features by making fuller use of data on quantities as well as prices (Berry 1994, Bresnahan 1981, Trajtenberg 1989). In order to estimate such a model, however, the methodology would first have to be extended to explicitly take into account the effects of network externalities and the resulting identification problem for the underlying demand curves.[26]

[26] Funding for this research was provided by the MIT Center for Coordination Science and the MIT Center for Information Systems Research. This work was made possible, in part, by the generous support of DataQuest, Inc., International Data Corporation and the National Software Testing Lab. Research assistance from D. Mayo, D. Young and Y. Kuo is gratefully acknowledged. Helpful comments were received from participants at research seminars at Boston University, Carnegie Mellon University, Copenhagen Business School, Georgia State University, MIT, the Naval Post-graduate School, Stanford University, University of California-Irvine, the University of Southern California, as well as the Fourteenth International Conference on Information Systems and the Fifth Workshop on Information Systems and Economics. Valuable comments from N. Gandal, H. Mendelson, R. Pindyck and five anonymous referees are also gratefully acknowledged. Any remaining errors are the responsibility of the authors.

References

Arthur, W. B., "Competing Technologies: An Overview," in G. Dosi (Ed.), *Technical Change and Economic Theory*, Columbia University Press, New York, Chapter 26, 1988.

Banker, R. D. and R. J. Kauffman, "Strategic Contributions of Information Technology: An Empirical Study of ATM Networks," *Proc. Ninth International Conf. on Information Systems*, Minneapolis, MN, 1988.

Belsley, D. A., E. Kuh, and R. E. Welsch, *Regression Diagnostics*, John Wiley and Sons, New York, 1980.

Berndt, E., *The Practice of Econometrics, Classic and Contemporary*, Addison-Wesley, Reading, MA, 1991.

—— and Z. Griliches, "Price Indexes for Microcomputers: An Exploratory Study," *Conf. on Research in Income and Wealth Workshop on Price Measurements and Their Uses*, Washington, DC, March 22–23, 1990.

——, M. H. Showalter, and J. M. Wooldridge, "On the Sensitivity of Hedonic Price Indexes for Computers to the Choice of Functional Form," MIT mimeo, Cambridge, MA, July 17, 1990.

Berry, S. T., "Estimating Discrete Choice Models of Product Differentiation," *RAND J. Economics*, 25, 2 (Summer 1994), 242–262.

Bresnahan, T. F., "Departures from Marginal Cost Pricing in the American Automobile Industry," *J. Econometrics*, 17 (1981), 201–227.

Brynjolfsson, E. and C. F. Kemerer, "Network Externalities in Microcomputer Software: An Econometric Analysis of the Spreadsheet Market," Working Paper 265, MIT Center for Information Systems Research, November 1994.

Chevalier, J. and D. Scharfstein, "Liquidity Constraints and the Cyclical Behavior of Mark-ups," *American Economic Rev.*, 85, 2 (May 1995), 390.

Chow, G. C., "Technological Change and the Demand for Computers," *American Economic Rev.*, 57, 7 (December 1967), 1117–1130.

Coffee, P., "Excel 3.0 Sets Spreadsheet Standard," *PC Week*, 8, 10 (March 11, 1991).

Cole, R., Y. Chen, J. Barquin-Stolleman, E. Dulberger, N. Helvacian, and J. Hodge, "Quality Adjusted Price Indexes for Computer Processors and Selected Peripheral Equipment," *Survey of Current Business*, 66, 1 (January 1986), 41–50.

Cringely, R. X., *Accidental Empires*, Addison-Wesley, Reading, MA, 1992.

Cusumano, M. A., Y. Mylonadis, and R. Rosenbloom, "Strategic Maneuvering and Mass-Market Dynamic: The Triumph of VHS Over Beta," *Business History Rev.*, 1992.

David, P. A., "Clio and the Economics of QWERTY," *American Economic Rev.*, 75 (May 1985), 332–337.

Epple, D., "Hedonic Prices and Implicit Markets: Estimating Demand and Supply Functions of Differentiated Products," *J. Political Economy*, 95, 1 (January 1987), 59–81.

Farrell, J. and G. Saloner, "Standardization, Compatibility, and Innovation," *Rand J. Economics*, 16, 1 (1985), 442–455.

—— and ——, "Installed Base and Compatibility: Innovation, Product Preannouncements, and Predation," *American Economic Rev.*, 76, 5 (1986), 940–955.

Fichman, R. and C. Kemerer, "Adoption of Software Engineering Process Innovations: The Case of Object-Orientation," *Sloan Management Rev.*, 34, 2 (Winter 1993), 7–22.

Fisher, F. M. and K. Shell, "Taste and Quality Change in the Pure Theory of the True Cost-of-Living Index," in Z. Griliches (Ed.), *Price Indexes and Quality Change*, Harvard University Press, Cambridge, MA, 1971, 16–54.

Gandal, N., "Hedonic Price Indexes for Spreadsheets and an Empirical Test of the Network Externalities Hypothesis," *RAND J. Economics*, 25, 1 (Spring 1994), 160–170.

Gordon, R. J., *The Measurement of Durable Goods Prices*, University of Chicago Press (for NBER), Chicago, IL, 1993.

Greenstein, S., "Lock-in and Costs of Switching Mainframe Computers Vendors: What Do Buyers See?," University of Illinois Faculty Working Paper 91-0133, April 1993.

Gurbaxani, V. and H. Mendelson, "An Integrative Model of Information Systems Spending Growth," *Information Systems Research*, 1, 1 (March 1990), 23–46.

Hartman, R. S., "An Empirical Model of Product Design and Pricing Strategy," *International J. Industrial Organization*, 7 (1989), 419–436.

—— and D. J. Teece, "Product Emulation Strategies in the Presence of Reputation Effects and Network Externalities: Some Evidence from the Minicomputer Industry," *Economics of Innovation and New Technology*, 1 (1990), 157–182.

Katz, M. and C. Shapiro, "Network Externalities, Competition, and Compatibility," *American Economic Rev.*, 75, 3 (1985), 424–440.

Matutes, C. and P. Regibeau, "Mix and Match: Product Compatibility Without Network Externalities," *Rand J. Economics*, 19, 2 (Summer 1988), 221–234.

Mayo, D. and D. Young, *Product Strategies for Packaged Software: An Exploratory Analysis of the Spreadsheet Market*, unpublished MIT Sloan School of Management S.M. thesis, Cambridge, MA, 1993.

Michaels, R., "Hedonic Prices and the Structure of the Digital Computer Industry," *J. Industrial Economics*, 27, 3 (March 1979), 263–275.

Morris, C. R. and C. H. Ferguson, "How Architecture Wins Technology Wars," *Harvard Business Rev.*, March–April 1993, 86–96.

National Software Testing Laboratories, "Software Digest Ratings Reports," NSTL, Conshohocken, PA, spreadsheet issues from 1985 to 1992 (1985).

——, "Macintosh Buyers Alert," NSTL, Conshohocken, PA, 1, 8 (1988a).

——, "Macintosh Ratings Reports," NSTL, Conshohocken, PA, spreadsheet issues from 1988 to 1991, (1988b).

——, "Macintosh Buyers Alert," NSTL, Conshohocken, PA, 1, 9 (1988).

——, "Spreadsheet Programs," *Software Digest Ratings Report*, 5, 6 (June 1988).

——, "Spreadsheet Programs," *Software Digest Ratings Report*, 7, 2 (February 1990).

Rao, H. R. and B. D. Lynch, "Hedonic Price Analysis of Workstation Attributes," *Comm. ACM*, 36, 12 (December 1993), 95–102.

Rao, V., "The Rising Price of Husbands: A Hedonic Analysis of Dowry Increases in Rural India," *J. Political Economy*, 101, 4 (August 1993), 666–677.

Rebello, K., M. Lewyn, and E. I. Schwartz, "Did Microsoft Shut the Windows on its Competitors?," *Business Week*, 32 (September 28, 1992).

Rosen, S., "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *J. Political Economy*, 82, 1 (Jan-Feb 1974), 34–55.

Saloner, G., "Economic Issues in Computer Interface Standardization," *Economics of Innovation and New Technology*, 1 (1989), 135–156.

—— and A. Shepard, "Adoption of Technologies with Network Externalities: An Empirical Examination of the Adoption of Automatic Teller Machines," *Rand J. Economics*, 26, 3 (1995), 479–501.

Stoneman, P., "Standards and Compatibility," in *The Economic Analysis of Technology Policy*, Clarendon Press, Oxford, 1987, 126–132.

Trajtenberg, M., "The Welfare Analysis of Product Innovations, with an Application to Computerized Tomography Scanners," *J. Political Economy*, 97, 2 (1989), 444–479.

Triplett, J. E., "The Economic Interpretation of Hedonic Models," *Survey of Current Business*, 86, 1 (January 1986), 36–40.

Westland, J. C., "Congestion and Network Externalities in the Short Run Pricing of Information Systems Services," *Management Science*, 38, 7 (July 1992), 992–1009.

Whang, S., "Contracting for Software Development," *Management Science*, 38, 3 (March 1992), 307–324.

White, H., "A Heteroskedasticity-Consistent Covariance Matrix and a Direct Test for Heteroskedasticity," *Econometrica*, 48 1980.

Accepted by Abraham Seidmann; received December 20, 1993. This paper has been with the authors 1 year for 2 revisions.

References

Hedonic Prices and Implicit Markets: Estimating Demand and Supply Functions for Differentiated Products
Dennis Epple
The Journal of Political Economy, Vol. 95, No. 1. (Feb., 1987), pp. 59-80.
Stable URL:
http://links.jstor.org/sici?sici=0022-3808%28198702%2995%3A1%3C59%3AHPAIME%3E2.0.CO%3B2-M

Hedonic Prices and the Structure of the Digital Computer Industry
Robert Michaels
The Journal of Industrial Economics, Vol. 27, No. 3. (Mar., 1979), pp. 263-275.
Stable URL:
http://links.jstor.org/sici?sici=0022-1821%28197903%2927%3A3%3C263%3AHPATSO%3E2.0.CO%3B2-B

The Rising Price of Husbands: A Hedonic Analysis of Dowry Increases in Rural India
Vijayendra Rao
The Journal of Political Economy, Vol. 101, No. 4. (Aug., 1993), pp. 666-677.
Stable URL:
http://links.jstor.org/sici?sici=0022-3808%28199308%29101%3A4%3C666%3ATRPOHA%3E2.0.CO%3B2-2

LINKED CITATIONS

Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition
Sherwin Rosen
The Journal of Political Economy, Vol. 82, No. 1. (Jan. - Feb., 1974), pp. 34-55.
Stable URL:
http://links.jstor.org/sici?sici=0022-3808%28197401%2F02%2982%3A1%3C34%3AHPAIMP%3E2.0.CO%3B2-U

The Welfare Analysis of Product Innovations, with an Application to Computed Tomography Scanners
Manuel Trajtenberg
The Journal of Political Economy, Vol. 97, No. 2. (Apr., 1989), pp. 444-479.
Stable URL:
http://links.jstor.org/sici?sici=0022-3808%28198904%2997%3A2%3C444%3ATWAOPI%3E2.0.CO%3B2-N

Congestion and Network Externalities in the Short Run Pricing of Information System Services
J. Christopher Westland
Management Science, Vol. 38, No. 7. (Jul., 1992), pp. 992-1009.
Stable URL:
http://links.jstor.org/sici?sici=0025-1909%28199207%2938%3A7%3C992%3ACANEIT%3E2.0.CO%3B2-N

Contracting for Software Development
Seungjin Whang
Management Science, Vol. 38, No. 3. (Mar., 1992), pp. 307-324.
Stable URL:
http://links.jstor.org/sici?sici=0025-1909%28199203%2938%3A3%3C307%3ACFSD%3E2.0.CO%3B2-7