Forthcoming, RAND Journal of Economics

Hedonic price indexes for spreadsheets and an empirical test of the network externalities hypothesis

Neil Gandal*

* Tel Aviv University. I am especially grateful to James Poterba and Manuel Trajtenberg. I would also like to thank seminar participants at the NBER Summer Workshop on Productivity, Boston University, Harvard University, Hebrew University, NYU, Tel Aviv University, University of Alberta, University of Calgary, and the University of Pennsylvania for helpful suggestions.

Abstract

In this paper, I first estimate hedonic price equations for computer spreadsheet programs, and then use the analysis to empirically test whether network externalities exist in this industry. The study shows that consumers are willing to pay a significant premium for spreadsheets that are compatible with the Lotus platform and for spreadsheets that offer links to external data bases, and a smaller premium for spreadsheets that offer local area network compatibility. These results support the hypothesis that the computer spreadsheet market exhibits network externalities. Finally, the quality-adjusted (real) price index of computer spreadsheets declined by approximately 15 percent per year from 1986-1991.

1. Introduction

A network externality exists when the value of consuming a particular product or service increases in the number of consumers that use compatible products or services. Network externalities arise from compatibility among products because compatibility gives consumers access to other firms' networks and stimulates the market for complementary products.

In recent years there has been great deal of interest in issues related to compatibility, standards and network externalities. Although there has been some empirical work, the literature has been generally theoretical.[1] This paper is an empirical investigation of the computer spreadsheet market. The goals are to (1) test whether this market exhibits network externalities, (2) determine which features or characteristics are most highly valued by consumers and (3) construct quality-adjusted price indexes for computer spreadsheets.

The spreadsheet market provides an excellent environment to test for the presence of network externalities because:

(1) A de facto standard exists. The computer spreadsheet market for IBM personal computers and compatibles has been dominated by Lotus. Upon its introduction in 1983, Lotus 1-2-3 became the best selling spreadsheet in the market and eventually the de facto industry standard.[2] One key observable characteristic is whether a spreadsheet is compatible with the Lotus platform.

(2) Network externalities should exist in this market for two reasons. (i) Katz and Shapiro (1992) note that direct network externalities arise in the computer software market because consumers want to be able to transfer files among themselves. If spreadsheet programs are compatible, this greatly enhances the ability of consumers to share information. (ii) Since many of the best-selling spreadsheets are compatible with the industry standard, there is a critical mass to stimulate the market for peripherals. Network benefits arise from complementary products such as TSP (a statistical analysis program) and Paradox (a database management program) that are fully compatible with the Lotus format.[3]

If consumers place a significant value on compatibility, this would provide evidence of network externalities.[4] Since some of the spreadsheets are compatible with the Lotus platform, while

other spreadsheet programs are incompatible (see footnote 2), it is possible to test whether consumers place a premium on compatibility.

A second feature that measures the importance of network externalities is whether a spreadsheet provides links to external data bases. If such compatibility exists, the user has access to numerous external data networks. Additionally, some spreadsheets link users through a local area network. If consumers place a premium on spreadsheets that offer these two features, this would provide further support for the network externalities hypothesis.

In order to perform these tests, it is necessary to estimate a price (or willingness to pay) function. Since the data is both cross-sectional and time series with repeated observations on some models, a natural way to proceed is to estimate hedonic price equations. Assuming that satisfactory specifications can be obtained, one can then estimate quality-adjusted price indexes for spreadsheet programs and also determine the factors that are most highly valued by users, in addition to testing for network externalities.

The results from this study are that consumers are willing to pay a significant premium for spreadsheets that are compatible with the Lotus platform and for spreadsheets that offer links to external data bases, and a smaller premium for spreadsheets that offer local area network compatibility. These results support the hypothesis that the computer spreadsheet market exhibits network externalities.

The study also finds that with the exception of the "network" features, the qualities most valued by consumers are basic rather than sophisticated. Finally, quality-adjusted (real) prices of spreadsheets fell by approximately 15 percent per year from 1986-1991.

To my knowledge, there have been few empirical tests of the network externalities hypothesis. Saloner and Shepard (1990) test for the existence of network externalities in the adoption of automatic teller machines (ATM's). In particular, they test whether banks with a larger expected number of ATM locations will adopt the ATM technology sooner. Since expected network size is not an observable variable, they use the number of branches as a proxy. Their results indicate the presence of network externalities. In a study of the mainframe computer market, Greenstein (1992) finds that other things being equal, a firm with an IBM 1400 was as likely as any other firm to purchase an IBM when making a future purchase, while a firm with an IBM 360 was more likely to purchase an IBM than a firm that did not own an IBM 360. Since software for the IBM 1400 could not run on the following generations of IBM models

(360, 370, 3000, and 4300), while software for the IBM 360 could run on the 370, 3000 and 4300, his results can be interpreted as a demand for compatibility.[5]

There has not been any work on constructing hedonic price indexes for computer software, although software purchases account for a non-trivial portion of the expenses on a microcomputer system. On the other hand, there is a relatively large literature on quality-adjusted price indexes for computer hardware. In a study of the microcomputer industry, Berndt and Griliches (1993) find that on average, quality-adjusted prices for microcomputers have declined at a rate of approximately 28 percent per year over the 1982-1988 period. Gordon (1989) shows that quality-adjusted prices declined in the mainframe and mini computer market by approximately 22 percent per year over the period 1951-1984. Dulberger (1989) constructs quality-adjusted price indexes for computer processors and shows that prices have declined by approximately 18 percent per year over the period 1972-1984.

In Section 2, I discuss the data. In Section 3, the hedonic price equations are estimated, tests for network externalities are conducted, and price indexes are constructed. In section 4, the effect of compatibility with the Lotus platform is examined in greater detail. Section 5 provides concluding remarks.

2. The data

This data set was compiled from yearly reports issued by DATAPRO Research Group, from 1986-1991.[6] DATAPRO lists three separate categories of Spreadsheet Programs: (1) Standalone Spreadsheet Programs, (2) Integrated Spreadsheet Programs, and (3) Financial Planning (Modeling) Systems. Because the markets for the three categories are quite different, the sample used in this analysis is taken from the first category only, which is defined to be spreadsheet and graphics programs. The second category consists of programs that include spreadsheet and graphics capabilities, but also include word processing programs and database management programs, etc. The main differences between financial modeling programs and spreadsheet programs are that financial modeling programs are much more powerful than spreadsheets, but they require users to write their own applications. Spreadsheets on the other hand have well defined modeling formats which users can easily employ. An examination of the data makes it clear that the characteristics which are important in explaining consumer willingness to pay are quite different for these three categories.

Computer hardware and software are complementary products and the benefit from software consumption can only be realized if consumers purchase compatible hardware, i.e., hardware capable of running the particular software package. Berndt and Griliches (1993) have shown that dummy variables for hardware platforms are important characteristics in explaining variations in microcomputer prices. Thus, in order to focus exclusively on software effects, the sample was restricted to spreadsheet programs that were compatible with the IBM line of computers and IBM compatibles and used either PC-DOS, MS-DOS, WINDOWS, or OS/2 operating systems. For example, spreadsheets that were only compatible with the Macintosh platform were not included in the sample.[7]

Finally, there were only two products introduced before 1983 (when Lotus 1-2-3 was first introduced) which appeared in the sample and since these products only appeared once in the sample, they were removed in order to reduce the set of vintage dummy variables. Nothing in the analysis changes if these two observations are included.

This led to a sample (unbalanced panel) of 91 model-observations. In addition to the basic editing functions (such as copying or moving rows and columns, adding, subtracting, multiplying, and dividing numbers, and replicating algebraic formulas) common to all spreadsheets, the DATAPRO report contains the following characteristics and features. I briefly summarize the available data.

(1) The dummy variable LOCOMP takes on the value one if the program is compatible with the Lotus (WKS, WK1) format. Otherwise the variable takes on the value zero. If the program is compatible with the Lotus format, it is capable of exchanging files with Lotus and other spreadsheets that support the Lotus format.

(2) The dummy variable RECALC takes on the value one (zero) if the program can (can not) automatically recalculate when new entries are made.

(3) The dummy variable SORTING takes on the value one if the program can sort a group of data observations on at least two levels, and zero otherwise.

(4) The variable GRAPHS is a dummy variable that takes on the value one if the program is capable of performing all of the following basic graphs: pie, bar, and line. If the program cannot perform these basic functions, the variable GRAPHS takes on the value zero. These

basic functions are bundled because the early DATAPRO reports collected the data in this manner.

(5) The variable WINDOW takes on the value one if the maximum number of windows on-screen simultaneously is between two to fifteen and takes on the value two if this maximum is sixteen or more. If this feature is not available, the variable WINDOW takes on the value zero. This variable was defined in this manner because some spreadsheets offer a maximum that is limited only by hardware features (such as memory).

(6) The dummy variable LINKING takes on the value one if values in several worksheets can be updated at the same time.

(7) The dummy variable EXTDAT takes on the value one if the program provides links to external data bases, and zero otherwise. This link can be either proprietary or through SQL support. If this feature is available, databases on mainframes can be downloaded directly into the spreadsheet.

(8) Macros allow a user to automate repetitive tasks. There are essentially two ways to write macros in spreadsheets. The first way involves the use of programming-like statements (such as if-then-else). The dummy variable PROGRAM takes on the value one if macros can be written in this manner.

(9) Macros can also be written in "learn mode." In learn mode, the keystrokes that are to be replicated are typed and the spreadsheet converts these keystrokes into a Macro. The dummy variable LEARN takes on the value one if the program enables the users to automate repetitive tasks in this manner.

(10) The dummy variable LANCOM takes on the value one if the program has the capability of linking independent users through a local area network (LAN).

(11) The dummy variable PRINT takes on the value one if three or more of the following five advanced print functions are possible: Sideways printing, Background Printing, Preview Mode, PostScript Support, and Printing of non-contiguous worksheet portions.

(12) The variable PRESENT takes on the value one if worksheets and graphs can be printed on the same page OR if multiple printing fonts (and character sizes) are available. If both features

are available, this variable takes on the value two, while if neither feature is available, the variable takes on the value zero. Although it seemed natural to group the two presentation features together, nothing in the analysis changes if these features are entered as separate variables.

(13) The dummy variable LOTUS takes on the value one if the program is produced by Lotus Development Corporation and zero otherwise.

(14) The variable MINRC measures the power of the spreadsheet and is defined to be the minimum of the maximum number of rows and columns that the spreadsheet can handle.

(15) The variable PRICE is the list price for a single copy of the program.[8]

(16) The variable LPRICE is defined to be the natural log of the price.

(17) The variable LMINRC is defined to be the natural log of MINRC.

All of the above variables are well defined, avoiding the problem of subjective ratings. For example, the major trade journals make occasional product comparisons of spreadsheets available in a particular year and then provide an overall "report card grade" which ranks the programs, but this is a subjective rather than objective measure. Additionally, these product comparisons select five or six products, i.e., they do not include all the available spreadsheets in the market.

Careful searches of personal computer magazines made it clear that discount prices could be obtained only for a small fraction of the sample. Hence discount prices were not available for this study, nor were market share data.[9] Although, it would have been preferable to have such data, the available data is rich enough to accomplish the objectives of the study.

Variables (1) through (6) are considered to be basic spreadsheet features, while variables (7) through (12) are more advanced features. Some of these advanced features (EXTDAT, PRESENT, and PRINT) were not available until 1989. The following table (Table 1) contains descriptive characteristics for the seventeen variables defined above.

[Insert Table 1 here]

In addition, to the above information, the year in which the observation was taken and the date of introduction are available. This allows the construction of time, age and vintage dummy variables, which are important for the analysis. The time dummy variables are denoted TIME87, TIME88, TIME89, TIME90, and TIME91.[10]

Similar to the personal computer (hardware) market, most products in this market were less than two years old. In this sample, fully 54 percent of the spreadsheets were new, 26 percent had been available for one year, 9 percent were two years old and 11 percent had been available for three or more years.

Lotus was the dominant firm throughout the period. The other major firms in the spreadsheet market for IBM personal computers and compatibles from 1986-1991 include Microsoft (first with Multiplan and then with Excel), Computer Associates (various versions of SuperCalc), Paperback Software (VP-Planner), WordPerfect (PlanPerfect) and Borland (Quattro and Quattro Pro) in the latter part of the sample. While these firms accounted for 35 of the model-observations in the sample and much of the "high end" of the market, at least five relatively inexpensive products were available in each year throughout the sample period.

3. Hedonic price equations and tests for network externalities

The first column in table 2 shows a regression (#1) with all possible explanatory variables,[11] while the second column of the table contains a regression (#2) with the insignificant variables from the first regression omitted. Comparing these two regressions, the omission of insignificant variables has little effect on the estimates of the remaining variables, and the adjusted R^2 is slightly higher and the standard error is slightly lower for the second regression.[12]

[Insert Table 2 here]

These two regressions show that many of the important features in determining the price of the package are quite basic (GRAPHS, WINDOW, LOCOMP, LINKING) and have been provided by some of the packages for many years. Much of the real price decline that took place in 1988 was due to the significant increase in the number of spreadsheets in the market that (1) offered basic graphic capabilities, (2) were compatible with Lotus, and (3) offered windows or enhanced windows capabilities. See Figure 1.

[Insert Figure 1 here]

In recent years, many firms have enhanced printing capabilities, added presentation features, made the spreadsheet compatible with a local area network and achieved compatibility with external data bases. The first two regressions in Table 2 show that of the sophisticated features, the two advanced "compatibility" characteristics are the ones consumers value significantly, especially access to an external data base. There is also a Lotus "brand name" premium. If two spreadsheets are comparable in terms of characteristics, a Lotus spreadsheet will command a higher price.

In order to further evaluate the network externality hypothesis, and investigate whether parameters were changing over time, the sample was split into two subsamples, (1986-1988) and (1989-1991), and regression (#2) was estimated for each of these subsamples. The results from these regressions are reported in Gandal (1992) and show that with the exception of LANCOM, LMINRC, and LINKING, the parameters are stable over time.

Thus, the third regression in Table 2 includes the variables TLANCOM , TLMINRC, and TLINKING, which are respectively LANCOM, LMINRC, and LINKING interacted with a dummy variable for the second (1989-1991) sample period. The significance of TLANCOM in regression 3 provides evidence that the compatibility with a local area network is a highly valued feature. One interpretation is that the value of this type of compatibility has increased over time as users have become more familiar with the capabilities of local area networks.

The estimated coefficient for TLMINRC in regression 3 indicates that the value of additional worksheet size (as measured by an increase in the minimum number of rows and columns) decreased over time. This can be explained by the fact that the average value of MINRC in the earlier (1986-88) subsample was 157, while the average value of MINRC in the second subsample was 241. A natural interpretation is that an increase in spreadsheet power is important up to a point. Finally, the value placed on the LINKING function has declined over time.

The third regression is the preferred hedonic price equation for the full sample since it has a significantly higher adjusted R^2 and significantly lower standard error than the second regression. The fourth regression in Table 2 is the preferred regression equation fit to the new products in the sample. A comparison of regressions 3 and 4 shows that the characteristics that are significant for the full sample are also significant in the sample of new products.

Since Lotus compatibility and external data base compatibility are highly significant, and since local area network compatibility is relatively significant in all four regressions in Table 2, there is strong support for the hypothesis that the computer spreadsheet market exhibits network externalities.

3.1 Validity of hedonic price regressions

Berndt and Griliches (1993) discuss a specification test to evaluate the validity of hedonic price equations. Their criterion is that vintage dummy variables should be insignificant in a satisfactory hedonic price regression. I employ this test to show that the preferred regression (#3) of Table 2 is valid. The specification tests were conducted as follows. The unrestricted model including all the relevant characteristics and all possible time, age, and vintage dummy variables was estimated, yielding an unrestricted R^2 of .902. Fitting the restricted model (without the vintage dummy variables) yields an R^2 of .887. Employing the standard F test, there is strong evidence that the hedonic price equation is satisfactory (F_{calc} = 1.07, while the critical value is $F_{.05,8,65}$ = 2.08). The age dummies are also approximately zero. Testing the joint hypothesis that the age and vintage dummies are zero, yields F_{calc} = 1.27, while $F_{.05,11,65}$ = 1.94, which supports the hypothesis that the vintage and age effects are insignificant.[13]

3.2 Price indexes

Table 3 shows the quality-adjusted (hedonic) price indexes for regressions 2,3 and 4 from Table 2.[14] The numbers in parentheses are the percentage price declines from the previous year. The average yearly decline in quality adjusted prices was 15 percent for the full sample and 22 percent for new products in the sample.

[Insert Table 3 here]

4. Further tests of the significance of lotus compatibility

In the first subsample, 15 of the 40 model observations (38 percent) were not compatible with the Lotus platform and some of these "incompatible" spreadsheets were relatively expensive. This is not the case with the second subsample, in which 42 of the 51 model observations (82 percent) were compatible with the Lotus platform. Further, the spreadsheets not compatible with the Lotus platform in the second subsample ranged in price from $39.00 to $60.00, while the

spreadsheets that were compatible with the Lotus platform ranged in price from $60.00 to $695.00.

It is thus encouraging that the variable TLOCOMP (which is LOCOMP interacted with a dummy variable for the second (1989-1991) sample period) is insignificant in both regressions 3 and 4. Adding the variable TLOCOMP to regression 3, one obtains a coefficient (t-stat) of -0.12 (-0.53), while adding the same variable to regression 4 yields a coefficient (t-stat) of 0.02 (0.055). Thus, the Lotus compatibility parameter is essentially the same for both subsamples.

Despite the fact that some of the incompatible spreadsheets were relatively expensive in the (1986-1988) subsample, the average price of a lotus compatible spreadsheet in this subsample ($365.00) is much higher than the average price of an incompatible spreadsheet ($80.00). Thus this subsample was further restricted to spreadsheets that cost less than $200.00. In this restricted sample of non-premium spreadsheets, there were 24 observations, of which 9 were compatible with the Lotus platform. The mean price of lotus compatible spreadsheets in this subsample was $151.00 versus $80.00 for the 15 incompatible spreadsheets. Gandal (1992) shows that although the lotus compatibility effect declines slightly in magnitude, it is still highly significant. Thus, the effect of Lotus compatibility continues to hold for non-premium spreadsheets.

5. Conclusion

The cumulative evidence from the analysis in this paper provides strong empirical support for the hypothesis that network externalities exist in the computer spreadsheet market. This paper also shows that it is possible to find well defined characteristics that explain the variation in software prices and hence construct quality-adjusted prices for software.

One limitation of the study was that market share data were not available. Trajtenberg (1989, 1991) has developed a sophisticated method (based on the estimation of a discrete choice model) that uses market share data to construct price indexes and measure economic performance. His methodology explicitly incorporates the benefits due to new goods and increased variety.[15] Trajtenberg has shown that this procedure is far superior to the hedonic price approach during the initial stages of a product cycle and when the variety of product offerings increases. On the other hand, when the product is not changing substantially and when the variety of offerings is relatively constant, hedonic price indexes should perform reasonable well. The latter scenario seems to characterize the computer spreadsheet market from 1986-1991. Many of the

significant changes in product design occurred between 1978, when the first electronic spreadsheet program (Visicalc) was introduced, and 1983 when Lotus introduced 1-2-3. Although new features have been added, the product design of spreadsheets is virtually unchanged since 1983, and many of the characteristics that consumers value highly had been introduced by 1986. Additionally, the variety of offerings has not changed substantially from 1986 to 1991. Thus there are compelling reasons to believe that the hedonic method for constructing quality-adjusted price indexes performs reasonably well in this setting. Hopefully market share data will become available so that future work can incorporate the Trajtenberg approach.

The other limitation to the study was that discount prices were not available. An interesting question is whether the premium for Lotus compatibility would show up if discounted prices were used. The argument that the Lotus compatibility effect would not show up in the discount market relies on the assumption that more expensive spreadsheets have higher percentage discounts (relative to list price) than less expensive spreadsheets. If all spreadsheets were discounted at the same percentage rate, only the coefficient on the constant would change in a regression of discounted prices. It was possible to obtain discounted prices for a small subset (25 model-observations) of the entire sample. The range on discounts in this sample was from 30 to 50 percent. The average discount for premium spreadsheets (a list price of more than $200) was 40 percent off list price, while the average discount for inexpensive spreadsheets (a list price of less than $200) was 42 percent off list price, suggesting that the percentage discounts are approximately equal. Even if one made the extreme (and unrealistic) assumption that the discount market price of Lotus compatible spreadsheets was 50 percent of the list price, while the discount market price of incompatible spreadsheets was 30 percent less than list price and used this rule to generate "discounted" prices for the full sample, the Lotus compatibility effect would still be significant in the preferred regression (coefficient estimate of .32, t=2.53). This suggests that the effect of Lotus compatibility would be significant in a full sample of discount prices.

Appendix

Correlation Matrix: Significant Independent Variables.
Full Sample N=91.

	EXTDAT	GRAPHS	LANCOM	LINKING	LMINRC	LOCOMP	LOTUS
EXTDAT	1.00						
GRAPHS	.36	1.00					
LANCOM	.56	.33	1.00				
LINKING	.47	.23	.52	1.00			
LMINRC	.32	.62	.52	.40	1.00		
LOCOMP	.34	.66	.46	.27	.65	1.00	
LOTUS	.24	.29	.19	.06	.11	.23	1.00
WINDOW	.32	.47	.39	.43	.44	.40	.18

1986-1988 sample (N=40) .

	GRAPHS	LANCOM	LINKING	LMINRC	LOCOMP	LOTUS
GRAPHS	1.00					
LANCOM	.31	1.00				
LINKING	.21	.23	1.00			
LMINRC	.63	.54	.35	1.00		
LOCOMP	.65	.43	.18	.72	1.00	
LOTUS	.26	-.20	-.26	.14	.22	1.00
WINDOW	.61	.32	.39	.57	.53	.10

1989-1991 sample (N=51) .

	EXTDAT	GRAPHS	LANCOM	LINKING	LMINRC	LOCOMP	LOTUS
EXTDAT	1.00						
GRAPHS	.45	1.00					
LANCOM	.64	.28	1.00				
LINKING	.56	.18	.69	1.00			
LMINRC	.38	.59	.44	.37	1.00		
LOCOMP	.40	.64	.41	.27	.50	1.00	
LOTUS	.22	.28	.34	.19	.04	.21	1.00
WINDOW	.41	.33	.41	.45	.28	.26	.21

References

Berndt, E. and Z. Griliches, 1993, "Price Indexes for Microcomputers: An Exploratory Study," in Price Measurements and Their Uses, Studies in Income and Wealth , M.F. Foss, M.E. Manser, and A.H. Young editors, NBER: University of Chicago Press.

Berry, S., 1991, Discrete Choice Models of Oligopoly Product Differentiation, mimeo.

Chou, C. and O. Shy, 1991, Network Effects without Network Externalities, *International Journal of Industrial Organization*, 8, 259-270.

Church, J., and N. Gandal, 1993, "Complementary Network Externalities and Technological Adoption," *International Journal of Industrial Organization,* 11, 239-260.

David, P., 1985, "Clio and the Economics of QWERTY," *American Economic Review Papers and Proceedings*, 332-337.

Dulberger, E., 1989, "The Application of a Hedonic Model to a Quality-Adjusted Price Index for Computer Processors," in Technology and Capital Formation, Dale Jorgenson and Ralph Landau editors, MIT Press.

Gandal, N., 1992, "Hedonic Price Indexes for Spreadsheets and an Empirical Test of the Network Externalities Hypothesis," Tel Aviv Foerder Institute for Economic Research, Working Paper 18-92.

Gilbert, R., 1992, "Symposium on Compatibility: Incentives and Market Structure, " *Journal of Industrial Economics,* Vol. XL, No.1, 1-8.

Gordon, R., 1989, "The Postwar Evolution of Computer Prices," in Technology and Capital Formation, Dale Jorgenson and Ralph Landau editors, MIT Press.

Greenstein, S., 1992, "Did Installed Base Give an Incumbent any (Measurable) Advantages in Federal Computer Procurement," *RAND Journal of Economics*, 24, 19-39.

Griliches, Z., 1971, "Hedonic Price Indexes Revisited," in Price Indexes and Quality Change, edited by Zvi Griliches, Harvard University Press, Cambridge, MA .

Katz, M., and C. Shapiro, 1992, "Product Introduction with Network Externalities," *Journal of Industrial Economics,* Vol. XL, No.1, 55-84.

Saloner, G. and A. Shepard, 1990, "Adoption of Technologies with Network Externalities: An Empirical Examination of the Adoption of Automated Teller Machines," Stanford University mimeo.

Trajtenberg, M., 1989, "The Welfare Analysis of Product Innovations, with an Application to Computed Tomography Scanners," *Journal of Political Economy*, Vol. 97, 444- 479.

_____. 1991, "Quality-adjusted Price Indices and the Measurement of Economic Growth," in Technology and Productivity: The Challenge for Economic Policy, O.E.C.D., Paris.

Footnotes

[1]A survey of the recent theoretical literature is provided by Gilbert (1992).

[2]Not all spreadsheets are compatible with the Lotus platform. In 1986, fully 42 percent of the spreadsheets in this study were compatible with the Lotus platform, while in 1989, 88 percent of the spreadsheets in this study were compatible with the Lotus platform. By 1991, "only" 76 percent of the spreadsheets in the study were compatible with the Lotus platform. See Figure 1.

[3]See Chou and Shy (1990) and Church and Gandal (1993) for theoretical work on "complementary" network externalities.

[4]If the size of the spreadsheet market had remained constant over time, a significant premium on Lotus compatibility could also be measuring a "lock-in" effect due to switching costs. Since the spreadsheet market grew at such a fast rate from 1986-1991, switching cost effects are insignificant.

[5] Several case studies show support for the network externality hypothesis. See, for example, David (1985).

[6] For 1987-1991, the month that the DATAPRO report was issued varies from May to July. Most of the 1986 DATAPRO report was issued in December 1985, while some portions were issued in March 1986. Several issues of Personal Computing and Infoworld magazines were useful in cross-checking some of the information.

[7]Many spreadsheets are now available in different versions that support the major platforms. For example, there is a version of Microsoft Excel 3.0 for the Macintosh and a version for the Windows operating system.

[8]It is standard practice in the software industry that the program is not actually purchased, but licensed (or leased) forever and that the vendor's copyright must be respected.

[9]Some personal computer magazines measured speed by having each spreadsheet perform a particular task; again these data were only available for a few spreadsheets in the sample.

[10]For example, the dummy variable TIME88 takes on the value one if the year is 1988 and zero otherwise.

[11]Since the vintage and age variables were insignificant in every regression, they are omitted here. Also note that variable RECALC is excluded from the regression. This is because nearly every spreadsheet offers this feature.

[12]The correlation matrix for the significant independent variables (regression (#2)) is displayed in the Appendix.

[13]The second regression in Table 2 also satisfies the Berndt-Griliches criterion. See Gandal (1992) for these calculations.

[14]For the second regression, price indexes are calculated by taking the exponentiated estimated coefficients on the time dummy variables, with the coefficient on T86 normalized to zero. For regressions 3 and 4, the procedure is slightly more complicated. See Berndt and Griliches (1993) for details.

[15]See also Berry (1991), who shows how discrete choice models can be used to estimate supply and demand models in markets with product differentiation.

Tables:

TABLE 1. DESCRIPTIVE STATISTICS: FULL SAMPLE (N=91).

VARIABLE	MEAN	S.D.	MAX	MIN
PRICE	275.4	205.8	695	35
LPRICE	5.25	.94	6.54	3.55
MINRC	285.0	261.5	1024	25
LMINRC	5.30	.89	6.93	3.22
EXTDAT	.24	.43	1.00	0.00
GRAPHS	.65	.48	1.00	0.00
LANCOM	.51	.50	1.00	0.00
LINKING	.59	.49	1.00	0.00
LOCOMP	.74	.44	1.00	0.00
LOTUS	.13	.34	1.00	0.00
LEARN	.67	.47	1.00	0.00
PRESENT	.34	.70	2.00	0.00
PRINT	.19	.39	1.00	0.00
PROGRAM	.80	.40	1.00	0.00
RECALC	.96	.21	1.00	0.00
SORTING	.67	.47	1.00	0.00
WINDOW	.86	.66	2.00	0.00

TABLE 2: REGRESSION RESULTS (DEPENDENT VARIABLE IS LPRICE)

VARIABLE	Regression (#1) All independent Variables	Regression (#2) Significant Variables	Regression (#3) Preferred equation	Regression (#4) New Models Only
	coeff. (t-stat)	coeff. (t-stat)	coeff. (t-stat)	coeff. (t-stat)
CONSTANT	3.73 (10.92)	3.76 (12.31)	3.12 (9.50)	2.61 (4.91)
TIME87	-0.076 (-0.45)	-0.062 (-0.38)	-0.07 (-0.43)	-0.57 (-1.73)
TIME88	-0.42 (-2.37)	-0.44 (-2.67)	-0.45 (-3.03)	-0.94 (-2.91)
TIME89	-0.64 (-3.50)	-0.70 (-4.20)	0.92 (1.71)	1.63 (1.65)
TIME90	-0.74 (-4.13)	-0.79 (-4.90)	0.90 (1.67)	1.47 (1.51)
TIME91	-0.82 (-4.48)	-0.85 (-5.30)	0.85 (1.59)	1.54 (1.56)
LMINRC	0.11 (1.31)	0.11 (1.59)	0.26 (3.24)	0.41 (3.42)
LOTUS	0.59 (3.46)	0.56 (4.36)	0.46 (3.62)	0.44 (1.97)
GRAPHS	0.45 (2.94)	0.46 (3.51)	0.52 (4.18)	0.36 (2.03)
WINDOW	0.17 (2.16)	0.17 (2.14)	0.14 (1.92)	0.18 (1.71)
LOCOMP	0.76 (5.30)	0.72 (5.28)	0.66 (5.17)	0.70 (3.02)
EXTDAT	0.52 (3.10)	0.55 (4.05)	0.57 (3.93)	0.67 (3.16)
LANCOM	0.25 (1.62)	0.21 (1.65)		
LINKING	0.18 (1.51)	0.21 (1.91)	0.26 (2.00)	0.43 (2.22)
LEARN	0.03 (0.18)			
PROGRAM	0.13 (0.70)			
PRESENT	-0.08 (-0.54)			
PRINT	0.20 (0.86)			
SORTING	-0.21 (-1.27)			
TLANCOM			0.61 (3.28)	0.33 (1.20)
TLMINRC			-0.34 (-3.07)	-0.52 (-2.76)
TLINKING			-0.31 (-1.49)	-0.25 (-0.76)
No. Observations	91	91	91	49
S.E. of Regression	.391	.385	.356	.385
R^2	.862	.857	.881	.875
ADJ. R^2	.827	.833	.857	.819

Table 3. Price Indexes for Spreadsheets (1986=1.00)

Year	1986	1987	1988	1989	1990	1991
Regression (#2)	1.00	.94 (6.0)	.64 (31.7)	.49 (22.6)	.45 (8.6)	.42 (5.5)
Regression (#3)	1.00	.93 (7.0)	.64 (30.9)	.50 (21.9)	.48 (4.0)	.46 (4.2)
Regression (#4)	1.00	.57 (43.0)	.39 (31.6)	.31 (20.5)	.27 (12.9)	.28 (-3.7)

Figures

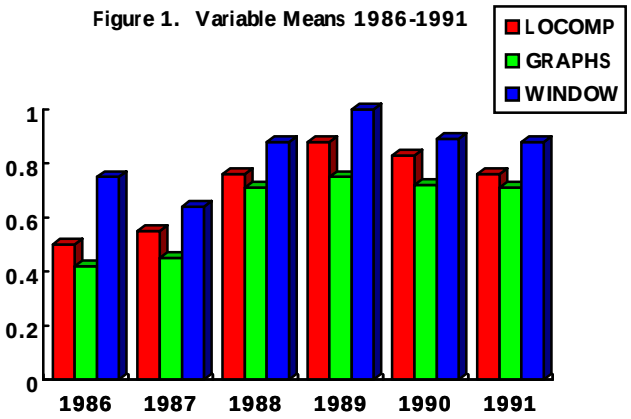
Figure 1. Variable Means 1986-1991
LOCOMP
GRAPHS
WINDOW
1
0.8
0.6
0.4
0.2
0
1986
1987
1988
1989
1990
1991