If the Software Fits: Capabilities, Ease of Use, and Cost of Econometric-Forecasting Microcomputer Software

by

Roger A. Dahlgran

IF THE SOFTWARE FITS: CAPABILITIES, EASE OF USE, AND COST OF ECONOMETRIC-FORECASTING MICROCOMPUTER SOFTWARE

Roger A. Dahlgran

Compared to spread sheet, word processing, data management, and statistical software, econometric software occupies a relatively obscure niche of the microcomputer software market. Nevertheless, this software is of great interest to economists and economic analysts because with the software and a microcomputer, the analyst can perform tasks that were difficult to perform with the large mainframe computers of a decade ago. Despite the power and flexibility of these programs, information about the specific capabilities of each is difficult to obtain because potential users have diverse professional interests as well as diverse technical requirements, both of which discourage econometric software vendors from advertising in a central medium. Consequently, objective comparative information about the software is not available. The objective of this report is to present a compilation of comparative information about econometric software for microcomputers.

The evaluation of econometric software must necessarily begin with a definition of its characteristics so that closely related software can be screened for further review. More specifically, though econometrics extensively utilizes multiple regression techniques, many of the 22 statistical packages reviewed by Carpenter, and many of the products offered by the 41 statistical software vendors listed in PC Week (March 17, 1987) are not well suited to estimating and manipulating econometric models. To establish a working definition, econometric software, as opposed to other types of software, is capable of performing a subset of the estimation and forecasting procedures discussed in standard intermediate econometrics texts, such as Pindyck and Rubinfeld, or Judge, et al. Judge, et al. present a two page comprehensive summary of the econometric models discussed in their text (pp. 750-1). Table 1 presents a further summarization of fundamental econometric model specifications. Therefore, econometric software should be capable of estimating, forecasting, and simulating a significant number of these models. Ten software products which are specifically targeted to performing econometric estimation, forecasting and simulation techniques were found. In alphabetical order, these products are: AREMOS, ESP, GAUSS, LIMDEP, MicroTSP, MODLER (also marketed as The Economists Workstation or EWS), RATS, SCA UTS, SHAZAM, AND SORITEC. A listing of current names and addresses of the vendors of these products is given in Appendix A.

Excluded from this evaluation are programs which emphasize forecasting at the expense of structural model estimation and analysis. These programs compute forecasts using exponential smoothing, and time series models but ignore estimation of structural econometric models. Examples of software in this class include 1-2-3-Forecast (a Lotus template), The Forecasting Edge, Smart Forecast II, Wisard Professional and Commercial Forecaster, and the AUTOCaST and 4CaST/2 products. Also excluded from this evaluation is SAS/PC, SPSS/PC+ and BMDP. These three programs are microcomputer adaptations of leading statistical-analysis mainframe programs, but they are not as well suited to the manipulation of econometric models as are the programs included in the evaluation. The mainframe version of SAS can be extended to include the Econometric and Time Series Library (ETS) which is a powerful econometric

---

Assistant Professor of Agricultural Economics, University of Arizona.

Table 1. Econometric estimation and forecasting problems to be addressed by econometric software for microcomputers.

Single Equation Structural Models

General Linear Model: (Fixed regressors, scalar identity covariance)
General Linear Model with Nonscalar Identity Covariance:
(Heteroscedasticity, autoregressive errors, other covariance structures)
Distributed Lag Models: (Geometric lags, polynomial lags)
Nonlinear Model
Qualitative Choice Models
Nonstable-parameter Models: (Random coefficients, switching regressions)
Restricted-parameter Models:
(Linear restrictions: exact, stochastic, inequalities. Baysian estimation)
Forecasting

Multiple Equation Structural Models

Disturbance Related Models
Parametric Restrictions Across Equations:
(Linear restrictions: exact, stochastic, inequalities. Baysian estimation)
Simultaneous Equations Models
Simulation

Time Series Models

ARIMA Models: (Identification, estimation, and forecasting)
Vector Autoregressive Models: (Identification, estimation, and forecasting)
Spectral Analysis: (Identification, estimation, and forecasting)

---

modeling program. However, the microcomputer SAS/ETS program is not scheduled for release until 1988. Also excluded were several microcomputer programs which contain Box-Jenkins time series estimation but not autoregressive and simultaneous equations estimation techniques. This category includes Minitab and SYSTAT.

GAUSS is ignored in comparing the features and capabilities of the software. This treatment of GAUSS is necessary because GAUSS is fundamentally different from the other programs. GAUSS is a matrix programming language which is useful for estimating and manipulating econometric models in matrix form. The other programs are command driven and estimate and manipulate econometric models in mnemonic form. Direct manipulation of the matrices is useful for specialized estimation problems, but for more usual econometric models, the canned procedures of the other programs are more comfortable. For these same reasons, the regression function of Lotus 1-2-3, which is likely used more than any of programs examined, is also not evaluated.

The accuracy of microcomputer econometric software is one criterion for evaluation. Most accuracy evaluations of regression software focus on the computational accuracy of the software in solving a nearly colinear model by ordinary least squares (see for example Malley and Monaco for an evaluation of

ESP, GAUSS, GIVE, LIMDEP, LOTUS, MicroTSP, RATS, SAS/STAT, SORITEC, and Y-STAT; and Eales for an evaluation of SHAZAM and RATS). Although ordinary least squares estimation is important in econometric work, other estimators, as well as forecasts and simulations, are also used frequently. The accuracy of these other algorithms is more difficult to document due to the use of subtly different procedures. In terms of numerical accuracy, the estimation algorithms of the econometrics programs generally perform well.

The other attributes by which econometric software can be judged are capabilities, ease of use, and cost. These three criteria are the main focus of this paper.

Capabilities

The capabilities of econometric software can be examined relative to minimum requirements, relative to the user's needs, and relative to other programs. First, the software selected for further evaluation has already met the minimum requirements for econometric software. Each program is capable of doing econometric analysis which frequently requires manipulating lagged data and estimating regression equations with serially correlated errors. Furthermore, the software performs the manipulations expressed algebraically in the econometrics texts and uses terms, concepts and references that are consistent with econometric theory.

In evaluating the software relative to the needs of the user, it is recognized that each economic researcher's software needs are unique. Specifically, the software needs of an academic econometrician, an academic researcher with commodity market interests, an academic research assistant, and a commodity market analyst in private industry are all different. The practicing econometrician will likely need direct matrix manipulation capabilities so that specialized estimators can be constructed, and will likely not be greatly interested in "canned" forecasting procedures. On the other hand, a commodity market analyst will likely have little interest in direct matrix manipulation, will need easy-to-use forecasting and simulation routines, and may need the ability to access commercial databases as well as the ability to easily generate presentation quality graphs.

Table 2 allows the individual user to compare the capabilities of each program to a list of the user's needed capabilities. More generally, Table 2 can also be used to compare the capabilities of the programs relative to each other. The cells of Table 2 are usually Y or N, indicating the availability or unavailability, respectively, of the indicated program feature. In some instances, the availability of features is not easily determined so the cell of the table is left blank, rather than to possibly misrepresent the product.

The table is divided into five major sections, each dealing with a general function of the software. The first section, entitled program operation, summarizes information about the software's processing modes, execution and data storage. The second and third sections of Table 2 respectively summarize the data manipulation and data management capabilities of each program. The fourth section shows the estimation and forecasting capabilities. The final section of Table 2 summarizes the graphics capabilities of each program.

Notice that the listing of the softwares' capabilities in Table 2 is much longer than the list of econometric specifications summarized in Table 1. In no way does this imply that any single program adequately addresses all of the specifications listed in Table 1. On the contrary, the bulk of Table 2 summarizes the programs' data manipulation and data management features. When

Table 2. Comparison of econometric software capabilities.

Program Version	AREMOS 2.0	ESP 2.0	LIMDEP Oct 86	MicroTSP 5.1	MODLER[a] 2.48	RATS 2.0	SCA-UTS 3.3	SHAZAM[b] 4.0	SORITEC 1.06b
Program operation									
Invisible copy protection		Y	Y	Y		Y		Y	N
RAM resident program	N	N	N	Y	N	Y			N
RAM resident accessible data	N	N	Y	Y	N	Y	Y		
Number of databases open	>1	1	0	0	<15	0	0	0	1
Online help	Y	Y	N	N	Y	Y	Y	Y	N
Interactive mode	Y	Y	Y	Y	Y	Y	Y	Y	Y
Command editor	Y	N	N	Y		N	N		Y
Prompting	Y	Some	N	Y	Y	N	Y		N
Menus	Y	Some	N	Y	Y	N	N		N
Batch mode	Y	Y	Y	Y		Y	Y	Y	Y
Resident program editor	N	N	Y	Y		N	N		N
Comment lines	N	Y	Y	N		Y	Y		Y
Program loops	Y	Y	N	Y		Y	Y	Y	Y
Conditional branching	Y	Y	N	Y		Y	Y		Y
Command replication	N	Y	N	Y		Y	Y		Y
Procedures/macros	Y	Y	Y	N	Y	N	Y	Y	Y
Variable lists	Y	Y	Y	N		N	Y		Y
Wildcard variable reference	Y	Y	Y	N		N	N		Y
Data Manipulation									
Transformation command	SERIES	GENR	CREATE	GENR	Not req'd	SET	Not req'd	GENR	Not req'd
Conditional manipulation		Batch only	Y	Y		Y	Y	Y	Y
Missing data handling		Y	Y	Y		Y	N	Y	Y
Addressable items									
Variables	Y	Y	Y	Y	Y	Y	Y	Y	Y
Variable groups	Y	Y	Y	N	N	N	Y		Y
Matrices	Y	Y	Y	N	N	Y	Y	Y	Y
Equations	Y	Y	N	Y	Y	Y	Y		Y
Procedures	Y	Y	Y	N	Y	N	Y		Y
Parameter estimates	Y	Y	Y	N	N	Y			Y
Cummulative density functions									
Normal	N	N	Y	Y		Y	Y	Y	Y
Student's t	N	N	Y	N		N	Y		Y
Chi square	N	N	Y	N		N	Y		Y
Fisher's F	N	N	Y	N		N	Y		Y
Other functions									
Lag function	x[-1]	x(-1)	x[-1]	x(-1)	x(-1)	x(T-1)	LAG(x)	LAG(x,n)	x(-1)
Difference function	Y	N	N	N	Y	Y	Y	N	Y
Probit function	N	N	Y	Y		N	Y		Y
Logit function	N	N	Y	Y		N	N		N
User defined functions	Y	N	N	N		N	N		N
Exponential smoothing	N	N	N	Y		N	Y		Y
Seasonal adjustment	Y	Y	N	Y	Y	N	Y		Y

Table 2. (Continued.)

Program Version	AREMOS 2.0	ESP 2.0	LIMDEP Oct 86	MicroTSP 5.1	MODLER[a] 2.48	RATS 2.0	SCA-UTS 3.3	SHAZAM[b] 4.0	SORITEC 1.06b
<u>Data Manipulation (Continued)</u>									
Frequency conversion									
Higher to lower	Y	Y	N	Y	Y	Y	Y		Y
Lower to higher	Y	Y	N	Y	Y	Y	N		Y
<u>Data Management</u>									
Data sources									
Console	Y	Y	Y	Y	Y	Y	Y	N	Y
ASCII file	Y	Y	Y	Y	Y	Y	Y	Y	Y
System data base	Y	Y	Y	Y	Y	Y	Y		Y
Comm. w/ commercial database	Y	N	N	N	Y	N	N		N
Data destinations									
Console	Y	Y	Y	Y	Y	Y	Y		Y
Printer	Y	Y	Y	Y	Y	Y	Y	Y	Y
ASCII file	Y	Y	Y	Y		Y	Y	Y	Y
System data base	Y	Y	Y	Y	Y	Y	Y		Y
Data editor	Y	Y	Y	Y	Y	Y	N		Y
Data labels	Y	N	N	N	Y	N	N		Y
List variables									
In memory/open database	Y	Y	Y	Y	Y	Y			Y
In unopened database	Y	Y	N	N	Y	N			Y
Rename variables									
In memory/open database	Y	Y	N	N	Y	Y			Y
In unopened database	Y	Y	N	N		N			N
Remove variables									
From memory/open database	Y	Y	Y	Y	Y	Y			Y
From unopened database	Y	Y	N	N		N			N
Copy variable									
Memory to database	Y	Y	Y	Y	Y	Y	Y		Y
Database to memory	Y	Y	N	Y	Y	Y	Y		Y
Database to database	Y	Y	N	N	Y	N	N		N
Sort data	N	N	Y	Y		Y	Y	Y	Y
DOS functions within program									
Directory	Y	Y	Y	Y	Y	N	Y		N
Rename	Y	Y	Y	Y	Y	N	Y		N
Erase	Y	Y	Y	Y	Y	N	Y		N
Copy	Y	Y	Y	N	Y	N	Y		N
View ASCII file	Y	N	Y	N	Y	N	Y		N
Print ASCII file	Y	N	Y	N	Y	N	Y		N

Table 2. (Continued.)

Program Version	AREMOS 2.0	ESP 2.0	LIMDEP Oct 86	MicroTSP 5.1	MODLER[a] 2.48	RATS 2.0	SCA-UTS 3.3	SHAZAM[b] 4.0	SORITEC 1.06b
<u>Estimation and forecasting</u>									
Single Equation Estimation									
Ordinary least squares	Y	Y	Y	Y	Y	Y	Y	Y	Y
Residual distribution test	N	N	N	N	N	N	N	Y	Y
Significance levels	N	N	Y	Y	N	N	N		Y
Autoregressive error models (order=p)									
Cochrane-Orcutt	N	p=1	p=1	$p \le P$	$p \le P$	$p \le P$	N	$p \le P$	1 or 2
Hildreth-Lu	N	p=1	p=1	N	$p \le P$	p=1	N	1 or 2	1 or 2
Maximum Likelihood	N	N	p=1	N	N	p=1	$p \le P$		N
Nonlinear estimation	Y		N				N		N
Lagged dependent variables	N	Y	Y	N		N	Y	Y	
Moving avg error (order=q)	N	N	Y	Y		N	$q \le Q$	$q \le Q$	
Generalized least squares	N	N	N	N		N	N	N	Y
Restricted least squares	N	N	Y	N		Y	N	Y	Y
Tests of restrictions	Y	N	Y	N		Y	N	Y	Y
Mixed estimation	N	N	Y	N		Y	N	N	Y
Almon lags	Y	Y	Y	Y	Y	Y	N	Y	Y
w/ autoregressive errors	N	Y	Y	Y		N	N		Y
Shiller lags	N	N	N	N		Y	N		Y
Box-Cox regressions	N	N	N	N		N	N	Y	N
w/ autoregresive errors	N	N	N	N		N	N	Y	N
Ridge regression	N	N	Y	N	Y	Y	Y	Y	Y
Nonlinear least squares	N	Y	N	Y		Y	N	Y	Y
Qualitative variables									
Logit	N	N	Y	Y		Y	N	Y	N
Probit	N	N	Y	Y		Y	N	Y	Y
Discrete choice	N	N	Y	N		N	N	N	N
Ordered probability	N	N	Y	N		N	N	N	N
Limited dependent variables									
Censored regression	N	N	Y	N		Y	N	Y	N
Truncated regression	N	N	Y	N		Y	N	N	N
Grouped data regression	N	N	Y	N		Y	N	N	N
Lognormal regression	N	N	Y	N		N	N	N	N
Switching regression	Y	N	Y	N		N	N	N	N
Sample selection models	N	N	Y	N		N	N	N	N
Random coefficients	N	Y	N	N		Y	N	N	N
Two stage least squares	Y	Y	Y	Y	Y	Y	N	Y	Y
w/ autoregresive errors	Y	Y	Y	N		Y	N	N	Y
Principal components	N	Y	Y	N		N	N	Y	Y
Single Equation Forecasting/Simulation									
Static	Y	Y	N	Y		Y	Y	Y	Y
Dynamic	Y	Y	N	Y		Y	Y	N	Y
Stochastic	N	N	N	N		Y	Y	N	N
Multiplier analysis	N	N	N	N		Y	N	N	N
Normalization on dep var		N	N	N	Y	N	N	N	N

Table 2. (Continued.)

Program Version	AREMOS 2.0	ESP 2.0	LIMDEP Oct 86	MicroTSP 5.1	MODLER[a] 2.48	RATS 2.0	SCA-UTS 3.3	SHAZAM[b] 4.0	SORITEC 1.06b
<u>Estimation and forecasting (Continued)</u>									
Multiequation Structural Models									
Linear SUR	N	Y	Y	Y	Y(Blue)	Y	Y	Y	Y
Restricted SUR	N	Y	Y	Y		Y	Y	Y	N
Mixed estimation	N	Y	N				N	N	N
Tests of restrictions	N	Y	N			Y	N	Y	N
Linear simultaneous equations									
Three stage least squares	N	Y	Y	Y	Y(Blue)	Y	Y(FIML)	Y	Y
Restricted 3SLS	N	Y	Y			Y	Y(FIML)	Y	N
Mixed estimation	N	Y	N				N	N	N
Tests of restrictions	N	Y	N			Y	N	Y	N
Nonlinear SUR	N	N	N	Y		N	N	Y	Y
Nonlinear simultaneous eqns	N	N	N	Y		N	N		Y
Multiequation Simulation									
Static	Y	Y	N	Y		Y	Y	N	Y
Dynamic	Y	Y	N	Y	Y	Y	Y	N	Y
Stochastic	N	N	N	N		Y	Y	N	N
Multiplier analysis	N	N	N	N		Y	N	N	N
Normalization on dep var		N	N	N	Y	N	N	N	N
Single Equation Time Series Models									
ARIMA models	N	Y	N	Y		Y	Y	Y	Y
Dynamic ARIMA models	N	N	N	Y		Y	Y	Y	Y
Spectral analysis	N	N	N	N		Y	N	N	N
Forecasting	N	Y	N	Y		Y	Y		N
Multiequation Time Series Models									
Vector autoregression models	N	N	N			Y	Y	Y	N
Spectral analysis	N	N	N			Y	N		N
Sim. trans. functions models	N	N	N			N	Y		N
Forecasting	N	N	N			Y	Y		N
<u>Graphics</u>									
Character plot	Y	Y	Y	N		Y	Y	Y	Y
High Resolution Graphics									
Line graph	Y	Y	N	Y	Y	Y	N		N
Scattergram	Y	Y	N	Y		N	N		N
Single range bar graph	Y	N	N	N		N	N		N
Multirange bar graph	Y	N	N	N		N	N		N
Stacked-bar graph	Y	N	N	N		N	N		N
Pie Chart	Y	N	N	N		N	N		N
Plotter support	Y		N	N	Y	Y	N		N

a/ Lack of detail due to general nature of the product's User's Manual.

b/ Lack of detail due to outdated Reference Manual.

doing econometric work, these features are as important as the ability to estimate the econometric specifications listed in Table 1 as the data must be properly transformed prior to submitting it to the estimation procedure.

Comparing just the estimation and forecasting section of Table 2 to Table 1, reveals that the software contains many specific estimation techniques, which are applicable to a single econometric specification. For example, a single equation autoregressive error model can be estimated by either the Cochrane-Orcutt procedure with or without the Prais-Winsten correction for the first observation and with or without iteration; by the Hildreth-Lu procedure; by the Durbin procedure; or by a Marquardt nonlinear estimation procedure. This single example illustrates that the breadth of econometric theory makes it unrealistic to expect any program to contain all possible estimation procedures. The programs are therefore better examined in terms of the user's need for specific algorithms and the cost or inconvenience imposed by not having access to those algorithms.

The cost of not having access to certain estimation procedures is not equal for all procedures. Most obviously, the cost of not having unneeded estimation procedures is zero. Beyond this, some estimation procedures can be performed by simply transforming the data prior to applying ordinary least squares. For example, with appropriate data transformations, autoregressive error models can be estimated with ordinary least squares. However, these models occur frequently and the data transformations are cumbersome so that a procedure to estimate autoregressive error models should be part of the econometric software. Less frequently needed estimators can be constructed by using the programming features of some software. For example, RATS and LIMDEP don't have a ridge regression procedure, per se, but the manuals show how to construct the estimator using the programs' features. Similarly, other estimators may be constructed with the clever use of the transformation and data manipulation facilities of the programs. In general, the greater the difficulty of performing the required data transformations and rearrangements, the greater is the inconvenience cost of not having the specific algorithm.

Several of the program features listed in Table 2 deserve special mention and clarification. Installable copy protection is the capability of the program to run from a required or available hard disk without having to insert a "key" disk into one of the floppy drives. This cumbersome copy protection method makes SORITEC more awkward to load than any of the other programs.

The RAM resident program and RAM resident data features relate to how the software operates. RAM resident programs initially load the entire program so that no additional program overlays are read from disk. As a result, RAM resident program execute faster.

As a group, the programs have two methods of accessing data. Some programs will access only the data in RAM (although virtual memory on disk may also be used) which can be read from, or periodically saved to, disk. Other programs access data in RAM as well as accessing several open databanks. Management of interrelated data sets is easier when several databanks can be open simultaneously.

Program operation is typically classified as batch and interactive. In interactive mode, the computer executes each instruction as it is received from the operator. Analysis in interactive mode is greatly facilitated by the presence of a command editor which edits recent commands so that typing mistakes can be corrected or alternative models estimated.

In batch mode, commands are read from a file of program statements. This mode is especially useful for performing complicated or lengthy procedures in a single run. Batch programming requires the use of an editor program to create and modify the batch programs. Memory resident utility programs, such as Sidekick, can edit batch programs. However if the econometric program requires 640K of RAM, no RAM is available for the utility program. Then it is desirable for the econometric program have its own batch program editor. A 640K RAM econometric program that does not have a batch program editor will be extremely difficult to use in batch mode because of the numerous program switches and reloads required to compose and debug the batch program.

In the section on data manipulation features, the most effective matrix manipulation capabilities are those allowed in conjunction with access to parameter estimates. Under these conditions, the user can enhance the set of estimators available in the program by combining existing procedures with specialized data matrices to generate sophisticated estimators applicable to specific problems.

Access to cumulative density functions is useful for hypothesis testing and significance-level reporting. Programs that neither report significance levels for F and t statistics, nor calculate cumulative densities, require the user to deal with the uncertainty about the significance level of the estimation statistics.

Some of the functions listed in Table 2 are not essential. For example, the logit transformation of a variable bounded by zero and one can easily be computed as the log of the odds ratio. Likewise, differencing is easily performed if lagged variables can be referenced. In contrast, other transformations, such as the probit transformation, are very difficult to perform if the function is not available as part of the software.

As the complexity of analyses increases, the data management capabilities of the programs become more important. Especially useful features are those for finding data buried in long forgotten data banks. These features include the ability to view DOS directories using DOS wildcard conventions, the ability to view the contents of a database without having to load the database to RAM, and the ability to attach labels to variables. Another useful feature is the ability to view an ASCII file, especially if it is not being read as expected.

An examination of the estimation, forecasting and simulation section of Table 2 reveals that some programs have strengths in estimating certain types of models. As their names suggest, LIMDEP (<u>LIM</u>ited <u>DEP</u>endent) is strong in the qualitative variables and limited dependent variables areas, and RATS (<u>R</u>egression <u>A</u>nalysis of <u>T</u>ime <u>S</u>eries) has strengths in time series models. Other programs, for example SORITEC, and MicroTSP cover the spectrum of estimation procedures with nonlinear simultaneous equations estimation algorithms at the top end of their spectra.

The strengths of other programs appears when their graphic capabilities are examined. For example, MODLER and AREMOS have only modest estimation capabilities but have comprehensive graphics capabilities.

In summary, all of the econometrics programs will read data from an ASCII file, will estimate regressions under several econometric assumptions and will save the data in a form that is suitable for subsequent analysis. A variety of other capabilities is also available. The programs can easily construct forecasts and perform simulations using the estimated model. The capabilities of the programs to handle time series data; to estimate models under assumptions that are specific to econometric theory; and to provide an interface between estimation, forecasting, and simulation is the source of utility for this software. However, this utility would largely be wasted if the software was not easy to use. The softwares' ease of use will be examined next.

Ease of Use

Table 2 demonstrates that no program clearly dominates the others in estimation capabilities. In fact, all of the programs can solve a wide variety of common estimation problems. In this situation, ease of use may serve as another evaluation criterion. The redefinition of ease of use to required operational effort makes the evaluation of the software's ease of use less abstract because required operational effort can be broken into two dimensions, qualitative effort and quantitative effort. Qualitative effort largely reflects the memorization and recall required to use the software. This effort is minimized with software that consistently utilizes econometric nomenclature in the command set so that the command required by the program to perform a desired operation is logically connected to the desired operation. The terms, definitions and procedures of econometric theory form a basis for this logical connection. Qualitative effort also includes the syntactical memorization-recall requirements for each command.

Quantitative operational effort is simply the keystrokes required to issue a given instruction. The trade off between qualitative and quantitative effort is apparent from an instruction like "PERFORM LEAST SQUARES ON DEPENDENT VARIABLE Y USING INDEPENDENT VARIABLES X1, X2, X3, X4 AND USE THE COCHRANE-ORCUTT PROCEDURE TO CORRECT FOR FIRST ORDER SERIAL CORRELATION WITH THE PRAIS-WINSTEN CORRECTION FOR THE FIRST OBSERVATION" which is efficient qualitatively but inefficient quantitatively. Although none of the programs evaluated approach this extreme, the nonabbriavated commands of SCA's UTS product are quite long. Easy to use programs are those that jointly minimize the required qualitative and quantitative effort.

Table 3 displays some pertinent points about the quantitative and qualitative effort required to operate three of the econometric programs reviewed. This table shows program segments required to perform a standardized analysis, which is to estimate a small dairy supply model as discussed by Dahlgran. The operations identified in Table 3 are

1. Initially allocate memory as required by the software.
2. Read data from the ASCII file DAIRYDAT.PRN.
3. Compute the natural logarithm of the price of dairy cows relative to the index of prices received by farmers, <u>i.e.</u> $\ln(PCW_t/PRC_t)$.
4. Compute a three year moving average of PCW_t/PRC_t.
5. Estimate the model $\ln(PCW/PRC)_t = \alpha_0 + \alpha_1 \ln(PAV/PFD)_{t-1} + \upsilon_t$ where PAV is the average price received by farmers for all milk, and PFD is the price of dairy ration.
6. Obtain the fitted values of the dependent variable, $\widehat{\ln(PCW/PRC)}_t$, from the above regression.
7. Use the fitted values to estimate the model

$$\ln(PCW/PRC)_t = \beta_0 + \beta_1 \widehat{\ln(PCW/PRC)}_{t-1} + \beta_2 \ln(PAV/PFD)_{t-1} + \beta_3 t + \varepsilon_t$$

where $\varepsilon_t = \rho\, \varepsilon_{t-1} + \nu_t$.

Items 5, 6, and 7 are the standard procedure for estimating a model with a lagged dependent variable and serially correlated errors.

An overall visual examination of Table 3 readily reveals differences in the required quantitative effort. Although the program segments accomplish the same tasks, some programs, notably LIMDEP and RATS in this comparison, require more keystrokes than do others, notably MicroTSP.

Table 3. Examples of program code required to accomplish a standardized task.

```
Operation    Program Code

        LIMDEP
 1&2. | READ; NREC=30; NVAR=21; FILE=B:DAIRYDAT.PRN;
      | FORMAT=(F2.0,F7.0,3F5.2,F5.1,2F6.0,3F5.0,F6.2,F5.1,F5.0,4F5.1,F5.0,2F6.0);
      | NAMES(X1=YR,X2=QT,X3=PAV,X4=PI,X5=PBL,X6=WI,X7=QFO,X8=QS,X9=PRC,X10=PPD,
      | X11=PWG,X12=PFD4,X13=PHY,X14=PCW,X15=PCL,X16=PSH,X17=PCV,X18=POP,X19=INC1,
      | X20=XCW,X21=XFD)$
      | .
   3. | CREATE; X=PCW/PRC; PCW=LOG(X)$
      | SAMPLE; 3-30$
      | CREATE;
   4. | X=PCW[-1]; PCW_MA3=PCW+X; X=PCW[-2]; PCW_MA3=PCW_MA3+X; PCW_MA3=PCW_MA3/3$
      | .
      | SAMPLE; 2-30$
   5. | CRMODEL; LHS=PCW; RHS=ONE,LRAVFD; KEEP=BL.PCW$
      | SAMPLE; 3-30$
   6. | CREATE; LBL.PCW=BL.PCW[-1]$
   7. | CRMODEL; LHS=PCW; RHS=ONE,LBL.PCW,LRAVFD,T; AR1; ALG=CORC$

        RATS
   1. | CAL 54 1 1; ALL 0 83:1;
   2. | OPEN DATA B:DAIRYDAT.PRN
      | DATA(UNIT=DATA,ORG=OBS, $
      | FORMAT='(F2.0,F7.0,3F5.2,F5.1,2F6.0,3F5.0,F6.2,F5.1,F5.0,4F5.1,F5.0,2F6.0)') $
      | 54,1 83,1 YR QT PAV PI PBL WI QFO QS PRC PPD PWG PFD4 PHY PCW PCL PSH PCV    $
      | POP INC1 XCW XFD
      | .
      | SMPL
   3. | SET PCW =LOG(PCW(T)/PRC(T))
   4. | SET PCW_MA3 =(PCW(T)+PCW(T-1)+PCW(T-2))/3
      | .
 5&6. | OLS PCW; # CONSTANT -RAVFD 1 1; PRJ BL_PCW
   7. | AR1 PCW; # CONSTANT -BL_PCW 1 1 -RAVFD 1 1 YR

        MicroTSP
   1. | CREATE A 54 83
   2. | READ(O) B:DAIRYDAT.PRN YR QT PAV PI PBL WI QFO QS PRC PPD PWG PFD4 PHY PCW PCL PSH PCV
      | POP INC1 XCW XFD
      | .
   3. | GENR PCW=LOG(PCW/PRC)
      | SMPL 56 83
   4. | GENR PCW_MA3=(PCW+PCW(-1)+PCW(-2))/3
      | .
      | SMPL 55 83
   5. | LS PCW C RAVFD(-1)
   6. | FIT BL.PCW
      | SMPL 57 83
   7. | LS PCW C BL.PCW(-1) RAVFD(-1) T AR(1)
```

When specific program segments are compared, differences in programming structures become apparent. For example, the LIMDEP read statement in Table 3 also renames the variables. In the absence of this renaming, LIMDEP default variable names are X1, X2, ... Though satisfactory for small problems, larger problems are easier analyzed if the renaming capabilities are used. Notice that the read statement for MicroTSP is the shortest because no format statement is used. Format statements are not allowed in MicroTSP which may create difficulties in reading some data sets. The RATS read statement requires several keywords and parameters.

Operations 3 and 4 indicate the ease of data transformations in the respective programs. Because LIMDEP allows only one transformation per expression, intermediate variables must be used. This requirement results in a lengthy statement for computing the three year moving average. RATS also uses an awkward transformation convention in requiring the time subscript "(T)" to be attached to each variable in each transformation. This requirement is noticeably cumbersome in operation number 4. Though not shown, SORITEC has the most efficient transformation convention in that neither transformation identifiers nor time subscripts are required.

The ordinary least squares commands, number 5, differ in their complexity primarily because of the syntactical requirements of RATS and LIMDEP. The corresponding statement in MicroTSP has fewer syntactical requirements. The four statements to accomplish operation 7 vary in their clarity. The AR(1) argument in the MicroTSP independent variables list signifies the estimation of a first degree autoregressive error process.

In summary, the measurement of qualitative operational effort is somewhat subjective. Nonetheless, for programs to be easy to use, command sets must be logically connected to the desired econometric operation and syntax requirements must be minimal. With all else equal, programs that don't minimize operational effort will not compete effectively in the software marketplace.

Cost

Having examined the capabilities and ease of use of the various econometric programs, the final evaluation criterion is cost. Table 4 summarizes the pertinent costs which are broken down as the direct costs of the software and the indirect costs of the required hardware. For programs with several available versions, the comparable statistics are given for each version. The different versions of each program are distinguished by the availability of features and the use of memory.

A LIMDEP spinoff, Econometrics Toolkit; MicroTSP; and SORITEC are all available in versions designed for ownership by students in econometrics courses. Although these programs are limited feature versions of the full-featured parent programs, they still contain enough features to allow students to perform realistic econometric analyses. For example, the student version of MicroTSP limits stored workfiles to 25 variables and 1,500 data points, and does not include linear and nonlinear system estimation. The SORITEC Sampler handles 1,500 data points, does ordinary least squares, two-stage least squares and autoregressive error models, comes with on-disk documentation, and duplication is openly encouraged. Either of these programs effectively remove computing cost concerns from teaching econometrics.

Other differences between the versions listed in Table 4 are due to the amount of memory used for data storage. For example, both RATS and SHAZAM have full-featured versions which employ different amounts of memory. Other differences between the versions reflect the features available. RATS and

Table 4. Program cost and hardware requirements of econometric microcomputer software.

Program - Version	Software: Price: one copy Academic ($)	Software: Price: one copy Other ($)	Software: Cost: ten copies Academic ($)	Software: Cost: ten copies Other ($)	Hardware: Hard Disk Req'd	Hardware: Hard Disk Space (Mbyte)	Hardware: RAM (DOS 2.0) Req'd (Kbyte)	Hardware: RAM (DOS 2.0) Rec'd (Kbyte)	Hardware: EGA Req'd	Hardware: 8087 Req'd
AREMOS										
Version 2.0	$3250	$3250	$20000+$5000/yr SL[a]		Y	3.0	640		N	Y
ESP										
Basic (w/o simulation)	716	795	1800	2000	N(1-2)[b]		256		N	N
Advanced (w/ simulation)	1166	1295	3600	4000	Y		512		N	Y
GAUSS, Version 1.49b										
With graphics	350	350	2250	2250	N(0)		256	320	Y	Y
W/o graphics	300	300			N(0)		256	320	N	Y
LIMDEP										
Oct 1986 version	250	250	500 SL	500 SL	N(1)		512		N	N
Econometrics Toolkit										
MicroTSP										
Version 5.1	295	595	1750	5950	N(0)		384	512	N	N
Student version of 5.0	40	N/A	2 free	N/A	N(0)		384		N	N
MODLER, Version 2.48 (EWS ver. 2.0)										
Single Eq'n	2025	2700			Y	2.7	640		Y	Y
25 Eq'n	2588	3450			Y	2.7	640		Y	Y
100 Eq'n	3150	4200	25130	33500	Y	2.7	640		Y	Y
1000 Eq'n	4650	6200	33750	45000	Y	2.7	640		Y	Y
RATS, Version 2.0										
Large memory	300	300	1650	1650	N(1)		512		N	N
Small memory w/ graphics	250	250			N(1)		256	256	N	N
Small memory w/o graphics	200	200			N(1)		256	256	N	N
SCA-PC System										
UTS Version 3.3	835	835	5010	5010	Y		512	640		Y
SHAZAM										
Small memory	250	250	1000 SL	1000 SL	Y	1.2	512	512	N	Y
Large memory	250	250	1000 SL	1000 SL	Y	1.2	640		N	Y
SORITEC										
Version 1.06b w/ support	195	345	1950	2950	Y	1.7	640		N	Y
Version 1.06b w/o support	195	195	1950	1950	Y	1.7	640		N	Y
Sampler (freely copyable)	25				N(0)		448		N	N

a/ SL designates site liscense.
b/ Number in parentheses designates the number of floppy disk swaps required to load program.

GAUSS have graphics and nongraphics versions, and ESP has a simulation and a nonsimulation version. The capabilities listed for ESP in Table 2, are based on the advanced version, which does simulation. Finally, UTS, is modular and its features depends on the modules purchased. Thus, the cost of the capabilities reported in Table 2 requires the reader to select the proper version from Table 4.

The first four columns of Table 4 reflect the academic and volume discounts that are frequently embedded in the pricing structure of microcomputer software. Prices for academic and other institutions are both reported to reflect the academic discounts. The cost of ten copies can be compared to the single copy price to approximate the slope of the volume discount structure. Site licenses are available for AREMOS, LIMDEP, SHAZAM, and may be advantageous for purchases of ten or more copies of an individual program.

One unique aspect of software pricing is the use of software leases. For example, Wharton Econometrics leases its software, AREMOS, for $3,000 for the first year. At the end of that time, the software may be purchased for $250. Sorities Group provides the user with the option of an annual lease or an outright purchase. The other vendors make their software available for outright purchase only. Leasing automatically entitles the user to support and updates, whereas the purchase of software exposes the user to additional upgrade charges as additional features are added to the program. Although these upgrade charges are usually minimal, inquiries about impending upgrades should be made before making a purchase commitment. Such an inquiry may result in a small saving if the purchase is postponed until the upgraded version is available.

SORITEC's pricing structure is unique in offering optional technical support. Note, however that support is included free of charge for the first sixty days on the purchase of the full featured SORITEC. The final feature of the pricing structure revealed in Table 4 is that the price of econometric software is highly variable. A casual comparison of the prices reported in Table 4 with the capabilities reported in Table 2 indicates no relationship between price and estimation capabilities.

The indirect cost of additional hardware required to run econometric software is a second cost component that must be considered. These costs may include required hard disk drives, additional memory and graphics boards, and mathematical coprocessor chips. The cost of upgrading a minimally configured machine may exceed the cost of the software itself. A rough cost estimate for user installed hard disks (for PC's not AT's) is $400 for a 20 Mbyte model. When hard disks are optional, the cost and convenience of the hard disk must be weighted against the inconvenience of floppy disk swaps. Random access memory (RAM) for a PC costs about $10 per 64 Kbytes, but the new RAM may require the purchase of an additional memory board ($140). Enhanced graphics adapter (EGA) cards cost from $300 to $500 and math coprocessor chips (8087) can be purchased for $125 for a PC and $200 for an AT. These cost estimates assume the hardware is purchased from the suppliers in personal computer magazines. If these items are purchased from a local computer store and installed there as well, the cost will be significantly higher.

Summary and Conclusions

In summary, the market for econometric microcomputer software was more populated than originally hypothesized. Microcomputer-based estimation algorithms exist in the econometric software for all the common and many of the uncommon estimation problems. Generally, the software is easy to use and as microcomputer capabilities increase the ease of use of individual programs should also improve. RATS, a notoriously difficult to use program, will soon be released for the Macintosh. Hopefully, the redesign for the Macintosh will focus attention on the user-computer interface, and result in increased ease of use for both the PC and the Macintosh version.

The pricing of the available programs covers a range that is almost as wide as the range of capabilities of the software. Pricewise, AREMOS and EWS have taken to the high ground while GAUSS, LIMDEP, MicroTSP, RATS, SHAZAM, and SORITEC are far less expensive. ESP and UTS occupy the middle ground. The capabilities and prices of the less expensive programs, make them look like bargains. In fact, an entire library of the less expensive programs will have far more capabilities than the expensive group and will cost less. The shortcoming of this library is that it will not be capable of accessing the Wharton database or model solutions. Also, the graphics capabilities of the inexpensive library will probably not be as good as the graphics capabilities of one of the more expensive products. Continuing improvements may narrow this gap.

Interestingly, while most of these programs are capable of computing forecasts from the estimated model, none of the software computes standard errors or prediction intervals for the forecasts. These computations are feasible given the computing power of the current generation of microcomputers. The continued omission of these computations amounts to another ambush that the econometric forecaster must watch out for. Evermindful that the forecast is simply an estimate, the forecaster will not know the quality of the estimate until the option of computing a prediction interval becomes available. Discussions between forecasters can then focus not only on the computed forecast but also the forecast's reliability.

References

Carpenter, E. H. "Statistical Packages for Microcomputers: New Vistas for Social Science Researchers in 1985." Social Science Microcomputer Review 3(1985):183-204.

Dahlgran, R. A. "A Synthesis of Microeconomic Duality Theory and Distributed Lag Modeling with Implications for U.S. Dairy Policy." North Central Journal of Agricultural Economics, 7(1985):132-44.

Eales, J. "Micro Statistical Computing." Computers in Agricultural Marketing, pp. 109-120. Madison: North Central Computer Institute, 1986.

Judge, G. G., R. C. Hill, W. E. Griffiths, H. Lutkepohl, T. Lee. Introduction to the Theory and Practice of Econometrics. New York: John Wiley and Sons, Inc. 1982.

Malley, J., and R. Monaco. "Numerical Accuracy in Econometric Analysis on Personal Computers." USDA/ARED/ERS working paper. Under review Journal of Computational Statistics and Data Analysis.

Minor, M. "Buyers Guide: Statistical/Analysis Software - Graphic Capabilities Offer Many Choices for Business Presentations." PC Week (March 17, 1987), pp. 55-65.

Pindyck, R. S., and D. L. Rubinfeld. Econometric Models and Economic Forecasts. 2nd. ed. New York: McGraw-Hill Book Company, Inc. 1981.

APPENDIX A. Econometric Microcomputer Software Vendors.

Package	Vendor
AREMOS/PC	Wharton Econometric Forecasting Associates 3624 Science Center Philadelphia, PA 19104
ESP	Economica Inc. 2067 Massachusetts Ave. Cambridge, Mass. 02140
GAUSS	APTECH Systems Inc. P.O. Box 6487 Kent, Wa 98064
LIMDEP	Department of Economics Graduate School of Bus Administration New York University 100 Trinity Place New York, NY 10006
MicroTSP	Quantitative Micro Software 4521 Campus Drive, Suite 336 Irvine, California 92715
Modler Econometric Software	Alphametrics Corp. 11 E. Princeton Road Bala Cynwyd, Pa. 19004
Economist Workstation	Data Resources/ McGraw Hill 24 Hartwell Ave. Lexington, Mass 02173
RATS	VAR Econometrics P.O. Box 1818 Evanston, Illinois 60204-1818
UTS	Scientific Computing Associates Lincoln Center, Suite 106 Lisle, Ill. 60532
SHAZAM	Kenneth J. White Department of Economics University of British Columbia Vancover, B.C. V6T1Y2 Canada
SORITEC	Sorites Group, Inc. 8136 Old Keene Mill Rd., A-309 P.O. Box 2939 Springfield, VA 22152-0939