software

SPSS/PC:
A Quantitative Historian's Dream or Nightmare?

Konrad H. Jarausch[C]

Anyone frustrated by the delays and breakdowns of mainframe computers, has experienced a recurring, seductive vision: If only the same powerful program package would run on one's own friendly microcomputer! In the last several months, this dream, born of anger over down-time and inter-minable turn-around, has come one step closer to reality: With much fanfare SPSS has introduced a microversion of its mainframe "Statistical Package for the Social Sciences" for the IBM PC. In the words of a promotional letter: "The overwhelming demand from users for a statistical package on an IBM PC has led us to the development of SPSS/PC."[1] Historians ought to have been curious about this announcement, since SPSS is the package most used by them on the mainframe. Moreover other stati-stical software for microcomputers has to this point fallen considerably short of their needs.[2]

Shoehorning a large and complex statistical package into a microcomputer was no mean feat. SPSS/PC therefore does have system and space require-ments, which will put it beyond the reach of most popular micros, used for word-processing (such as the Apple IIs or the Kaypro): It needs a IBM PC (or XT) with at least 320K of memory, a hard disk, also a math-coprocessor (an 8087 chip) to speed up its calculations, and MS DOS. The hard disk is necessary, since the program comes on nine (!) diskettes, each containing modules and subroutines, which need to be available concurrently. These are substantial hard-ware requirements, putting the necessary system into the top range of personal computers. But with the continued memory/size expansion of micros and dropping prices, such machine configurations are certainly within reach of departments and increasingly also of individual scholars.

In actual operation, the microversion of SPSS is remarkably similar to its bigger sister. It can handle up to 200 variables at the same time and disk space is the only limiter of the number of cases. In practice it runs relatively quickly for data sets of let's say 25 variables and 1000 cases, even if it has to switch modules along the way. The major syntax innovation is a period behind each command, since there is no longer a computer card ending which automatically indicate the command end. Should one forget to include the period, the package prompts the user to supply it or other additional information.[3] SPSS/PC must be run from the hard disk and cleverly requires a key diskette (as copy protection) to get it started each time. Fortunately, the program comes with a tutorial which

(')Address all communications to: Konrad H. Jarausch, The University of North Carolina at Chapel Hill, Department of History, Hamilton Hall, N.C. 27514, USA.

allows the novice to try out its main features before creating his own data set. In general the operation is, for such a large package on a small machine, efficient and user friendly. Especially for those scholars, already somewhat familiar with one of the mainframe versions, the transition should be fairly easy.

The initial working impression of SPSS/PC is therefore positive, not to say exhilerating. Finally, the historians can enter his own data, check them through a FREQUENCIES command and run simple procedures such as CROSSTABS (with the same statistics as on the mainframe) all in the same session at his desk in his own study! No need to punch cards any longer in a noisy and frantic user's room at the computer center or to fight for a terminal. This favorable experience is reinforced by the interactive nature of SPSS/PC which is decided step beyond the mainframe version: Not only are commands immediately displayed on the screen, but error messages also flash up in an instant so that mistakes can be corrected without frustrating waiting for a disappointingly thin printout of another "busted" run. Moreover the package has a help prompt which explains the basic features of any procedure so that the researcher can make required corrections without consulting printed instructions. The manual retains its accessibility for historians through a large "statistics guide" section which explains the purpose and limitations of different procedures on the basis of actual working examples. Moreover SPSS has kept many of its data-transformation capabilities (so important to historians). It also offers a broad array of descriptive and analytical statistics from simple data listing and FREQUENCIES all the way up to FACTOR and CLUSTER analysis. Finally, it has added a PLOT procedure (which will even run without a graphics card) and has entered the exciting world of "log-linear analysis," a promising statistical development for historians concerned with nominal and categorical data.(4) The resulting table display is, given the limitation of screen size, attractive and easily printed in the format desired. On the whole the microversion of SPSS is therefore closer to the new SPSS-X release for the mainframe than to the earlier version, and in its interactive behavior, even a step beyond.

Skeptics are likely to ask: This seems to good to be true, what is the catch? The first several weeks of working with SPSS/PC did, indeed, reveal a number of shortcomings. To begin with, some power had to be sacrificed in the "downsizing." Therefore the DISRCIMINAT, MANOVA and SURVIVAL procedures were dropped from the microversion.(5) Second, data files cannot created within SPSS, but must be produced by a text editor, such as the clumsy EDLIN (in DOS) or the fancier word processing packages like Microsoft word. One needs to remember, however, that SPSS will accept only standard ASCII files, a limitation which produces gibberish when not obeyed! Third, SPSS/PC will process merely one file at a time, which is a serious drawback for historians, since it means that they need to do their file merging, etc. in some other software environment or on the mainframe computer. However, SPSS now also utilizes the matrixes created by simply procedures like crosstabulations as data for higher statistics such as factor analysis, a considerable convenience. Finally, the asking price for the copy-protected package is $ 795, a rather steep cost for an individual scholar. Initially, the program was not sold by buying clubs, although one would expect that with proliferation some discounts will become available (for instance if a store orders more than 10 copies, the price drops to the mid-500s). While some of these draw-

backs (such as leaving out some procedures) are understandable, users *ought to be* warned more clearly about the need for a text editor. And one would hope that the single file limitation and the communications problems with the mainframe will be remedied in future releases.

Is SPSS/PC therefore the long-awaited dream or yet another software nightmare? The answer to this question depends somewhat upon expectations and alternatives. While a microversion of BMPD has apparently come out, this bio-medical package is rarely used by historians. The more frequently employed SAS has the problem of being written in PL1 which is not available on IBM PCs and is therefore going to take a considerable time in coming up with a micro-conversion. Compared to the earlier packages (ABSTAT), SPSS/PC wins hand down.(6) Its strength is with small to medium size data sets of one to two dozen variables. Those files SPSS/PC can handle quickly, efficiently and interactively. Really large data sets or hughe numbers of variables are still better processed on the mainframe. To facilitate such communication SPSS does, fortunately, also contain an up- and downloading routine, called KERMIT.

Since historians are more likely to work with moderate numbers of cases and variables, SPSS/PC marks a decided advance in their working tools. While a final evaluation will have to await the experience of hundreds of hours of running-time of all the intricate operation, data-definition and procedures commands, the initial impression of SPSS/PC is quite favorable. The microversion of this statistical package has the great advantage of disengaging quantitative methods from the big-technology environment and for putting great statistical powers into the hands of the individual researchers. In many way SPSS/PC therefore promises to be more compatible with the artisanal work-style of historical craftsmanship than its mainframe counterpart.(7)

NOTES

1 "SPSS/PC and SPSS/Pro Open SPSS to Micro World," Keywords, 33 (1983), 2ff. It was actually first available on the DEC Professional 500.
2 D.K. Rowney, "The Historian and the Microcomputer," Byte, 7 (1982), 166-176; R. Jensen, "The Microcomputer Revolution for Historians," Journal of American History, 14 (1983), 91-111; and M.H. Finefrock, "The Hard Software Decision," (part II of "Computers and the Historian"), AHA Perspectives, 21 (1983).
3 M.J. Norusis, SPSS/PC for the IBM PC/XT (Chicago, 1984).
4 J.M. Kousser and others, "Log-Linear Analysis of Contingency Tables: An Introduction for Historians," Social Science Working Paper, 417 (1982) of the California Institute for Technology.
5 N. Nie and others, SPSS: Statistical Package for the Social Scienes (Chicago, 1975) 2nd ed. and SPSS-X Users" Guide (Chicago, 1983).
6 R. McCaa, "Microcomputer Software Designs for Historians," Historical Methods, 17 (1984), 68-74 fn. 7 lists the relevant literature.
7 K.H. Jarausch, G. Arminger and M. Thaller, Quantitative Methoden in der Geschichtswissenschaft: Eine Einführung in die Forschung, Datenverarbeitung und Statistik (Darmstadt, 1984).